HEPATITIS B CORE EVOLUTION AND CELLULAR LOCALISATION

BY

SEYED MOHAMMED JAZAYERI

A thesis presented for the degree of Doctor of Philosophy,

IBLS

University of Glasgow

Division of Virology

Church Street

Glasgow

G11 5JR

October 2003

To my dear wife and daughters,

Reihaneh and Sepideh

that really without whose affection, love, support and patience,

this thesis would not have been possible

Finally, last but not least, without the support of my wife I would never have started the work towards this thesis. Also, I wish to thank my parents for relentlessly begging the question "Have you finished it yet?"

-The author was in receipt of Islamic Republic of Iran student scholarship. I thank The Ministry of Health and Medical Education for generous financial support. Unless otherwise stated, all work presented in this thesis was carried out by the author.

Seyed Mohammed Jazayeri **October 2003**

Summary

HBV core gene has attracted a resurgence of interest recently because it influences the outcome of hepatitis B infection. The aims of this study were: to characterise the core gene variability in diverse geographic regions and ethnic groups; to determine subtype/genotype-specific variants; to specify country/ethnic-specific variants; and to identify any relationship between intracellular HBcAg and chronic HB disease.

The current work can be divided into three main parts. In chapter 3, in the first and second sections, HBV DNA in 139 sera samples from 12 diverse geographic origins was analyzed by sequencing the core gene. We found that certain residues allow definition of subtypes and genotypes. Further, specific nucleotide motifs were defined for particular countries. We also found, intriguingly, a set of amino acid variants in a majority of sequences from South-east Asian patients and also for Western populations. In the second section of chapter 3, core gene variability in the Pacific region, the nucleotide diversity of the genotype C samples was significantly greater than that of genotype D samples, which is consistent with a longer history of HBV infection in genotype C predominant islands. Results obtained from this chapter confirm that HBV strains spread within constrained ethnic groups and that selection pressure defined the sequence variability within each genotype (and subtype). It may be that specific T cell epitopes are specific for particular geographic regions, and thus ethnic groups, which may affect the design of immunomodulatory therapies, including therapeutic vaccines.

The third section of this thesis (chapter 4) deals with in vitro biological effects of HBV C gene variability. Confocal microscopy was used to visualise HBcAg cellular expression in 40 clones (including mutagenised samples) from various stages of HB disease from

patients with different HBeAg/anti-HBe status.. This was correlated with sequence variation within B cell epitopes and at the C-terminus. Using site-directed mutagenesis to revert C-terminus and B-cell epitope mutated sequences with cytoplasmic expression to the original (wild type) sequences led to a shift back to predominantly nuclear and both nuclear and cytoplasmic distribution, respectively.

In conclusion, we showed that HBcAg reflects the genotype of the whole genome and there are certain residues that are distinct for the HBV genotypes. Moreover, we demonstrate that the pattern of HBcAg localisation in vitro is dependent on sequence and correlates with the serology of chronic HBV infection. It supports the hypotheses that such variants may play a biological role including antigenic expression. This study not only has an impact on the design of the therapeutic and prophylactic vaccines, but on our understanding of the immune response to different HBVs.

Chapter 5 is composed of a preliminary study on FTO9.1 cell line infectivity by HBV. This cell line which is transfected with a construct containing human-Annexin V, showed propagation of virus in primary and transfer experiments. Due to some difficulties using the Cat III laboratory, this project stopped after the first six months of the research course. The results obtained during that period have therefore been included in a separate chapter (chapter 5).

Table of contents

ABBREVIATIONS

aa	Amino acid
Ab	Antibody
Ag	Antigen
ALT	Alanine aminotransferase
bp	Base pairs
CAH	Chronic active hepatitis
ccc	Covalently-closed circular
C-terminus	Carboxy terminal end of protein
CTL	Cytotoxic T-lymphocyte
dATP	2'-deoxyadenosin 5'-triphosphate
dACT	2'-deoxycytidine 5'-triphosphate
dGTP	2'-deoxyguanosine 5'-triphosphate
dTTP	2'-deoxythymidine5'-triphosphate
DHBV	Duck hepatitis B virus
DNA	Deoxyribonucleic acid
dNTP	2'-deoxynucleoside 5'-triphosphate
EDTA	Ethylenediaminetetra-acetic acid (disodium salt)
ER	Endoplasmic reticulum
EtBr	Ethidium bromide
FITC	Fluorescein isothiocyante
h-A V	Human annexin V
HBcAg	Hepatitis B core antigen

HBeAg	Hepatitis B e-antigen
HBsAg	Hepatitis B surface antigen
HBV	Hepatitis B virus
HCC	Heparocellular carcinoma
HCMV	Human cytomegalovirus
HCV	Hepatitis C virus
HIV	Human immunodeficiency virus
HLA	Human leukocyte antigen
IFN	Interferon
IL	Interleukin
IPTG	IPTG (isopropyl-β-D) thiogalactopyranoside
mAb	Monoclonal antibody
MHC	Major histocompatability complex
MOPS	(3-(N-Morpholino) propanesulfonic acid (pH 7.0)
mRNA	Messenger RNA
NLS	Nuclear localisation signal
NPC	Nuclear pore complex
ORF	Open reading frame
PBS	Phosphate-buffered saline
PCR	Polymerase chain reaction
pgRNA	Pregenome RNA
RC	Relaxed circular
rhAV	Recombinant human annexin V

RNA	Ribonucleic acid
RNase	Ribonuclease
rpm	Revolutions per minute
SDS	Sodium dodecyl sulphate
Th	T-helper
Tween 20	Polyoxyethylene sorbitan monolaurate
uv	Ultraviolet

LIST OF FIGURES

LIST OF TABLES

CHAPTER 1 GENERAL INRODUCTION

1.1 HBV STRUCTURE AND GENOME ORGANIZATION

Hepadnaviruses are small hepatotropic DNA viruses that can produce persistent infections of liver cells, leading to chronic hepatitis and, in some cases, to hepatocellular carcinoma. The prototype hepadnavirus is the human hepatitis B virus (HBV), but related viruses have also been isolated from several other vertebrate species including squirrels, woodchucks, and ducks. Electron microscopy examination of partially purified serum derived preparations of HBV revealed the presence of three kinds of particles. Complete infectious virions of about 42-nm in diameter which consist of a 36-nm icosahedral nucleocapsid of the hepatitis B core antigen (HBcAg) and a 7nm lipoprotein bilayer derived from the endoplasmic reticulum (ER) membrane of the host (Dane et al., 1970; Patzer et al., 1986; Crowther et al., 1994). Three HBV surface proteins of varying sizes (see section 1.3.2) are inserted into this lipoprotein bilayer. The viral DNA, a virus encoded RNA-dependent DNA polymerase, a protein kinase which phosphorylates HBcAg (see section 1.3.1B) and a genome bound protein, covalently linked to the negative strand of HBV DNA, are all contained in the nucleocapsid (Albin and Robinson, 1980; Gerlich and Robinson, 1980).

Small spherical and filamentous forms of non-infectious subviral particles of 20-nm diameter are composed of S protein with variable amounts of M and L protein in addition to the host derived lipids (Ganem et al., 1996). Although they do not contain viral DNA or nucleocapsid, they are highly immunogenic and elicit a strong antibody response in most individuals. The concentration of subviral particles often exceeds the virion concentration by at least three fold.

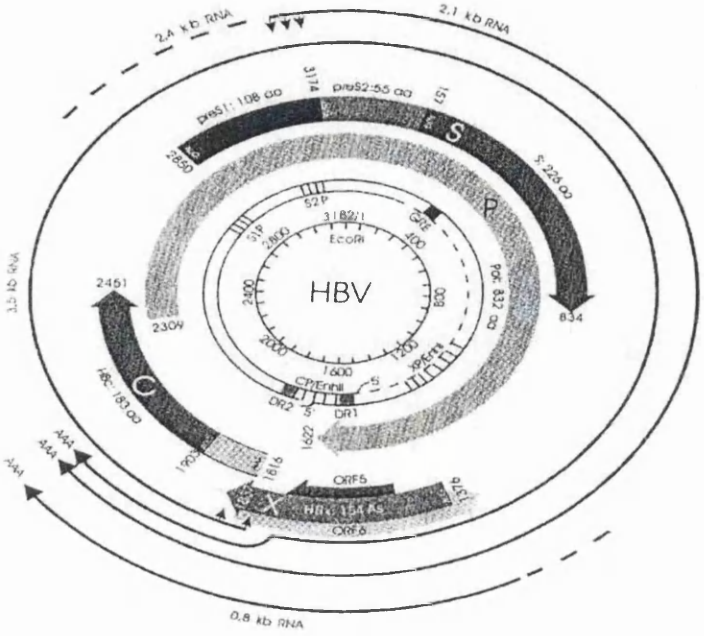


Figure 1.1: Genomic organisation of HBV.

The outer lines represent the different classes of transcripts. The four major ORFs (S, P, C and X) are indicated in the centre. The partially double stranded circle represents the viral genome, showing the promoters (S1P, S2P, XP and CP), the enhancers (EnhI and Enh II) and direct repeats (DR1 and DR2). Taken from Caselmann, 1996.

Within the core is the 3 to 3.3-kb viral genome and a polymerase activity known to be central to its replication. The virion DNA of these viruses is a relaxed circular species that is only partially duplex (Fig 1.1). Although the minus strand is full length, its plus strand complement is less than unit length, being variably truncated at its 3' end. The result is a single stranded gap of varying extent (20-80% of the genome). It has been shown that host DNA polymerase fills in this gap (reviewed by Kann et al., 1998). The most important features of the genomic DNA, however, are found at its termini: the minus DNA strand has a protein covalently attached to its 5' end, whereas the 5' end of the plus strand is covalently linked to a capped oligoribonucleotide. The coding organization of the viral DNA is also summarized in Fig 1.1. The coding organization is highly compact, i.e. every nucleotide is within a coding region, and over half of the sequence is translated in more than one frame. As shown in figure 1.1, there are four

major open reading frames (ORFs), all encoded by minus-strand DNA (Ganem et al., 1987), these are the surface (S), core (C), polymerase (P), and X protein ORFs. These proteins are translated from a series of viral transcripts whose extents are also summarized in fig.1.1 (see below).

1.2 HBV life cycle

Little is known about the early molecular events regarding the nature of the virus-receptor interaction, viral uptake or uncoating. However, considerably more is known about viral replication, virion release and later events in viral replication. The narrow host range of HBV and lack of suitable cell lines for infection are major limitations to such studies (for details see chapter 5).

1.2A Attachment to the cell membrane

To date the exact mechanism of attachment and penetration of HBV into human hepatocytes is not fully identified. Binding of a viral particle to the human hepatocyte plasma membrane is an important step in viral entry and replication of HBV. In this interaction, the envelope proteins of HBV are thought to play a crucial role. Although there have been many reports during the last two decades of putative receptors, the identity of the HBV receptor remains elusive.

It has been suggested that the pre-S1 domain, amino acid residues 21-47, is probably the most important attachment site to liver cells (De Meyer et al., 1997; reviewed by Cooper, 2003). Besides hepatocytes, many cells of non hepatic origin including hemopoetic cells of the B lymphocyte lineage, peripheral blood lymphocytes and even some non-human, simian virus 40 transformed cell lines have receptors for the pre-S1 (21 to 47) region (Budkowska et a l., 1993). Several receptors on the liver cell membrane have been proposed (Table 1.1). Viral pre-S1 interacts with several proteins, which have therefore been suggested as receptor candidates. Asialoglycoprotein receptor (ASGPR), a transmembrane molecule specifically

expressed on the hepatocellular membrane (Treichel et al., 1997), human interleukin-6 (Neurath et al., 1992), and transferrin (Franco et al., 1992), are all proposed to contain recognition sites for the HBV pre-S1 domain. Also, peptides of molecular weight of 31 kd (Dash et al., 1992), which interact with pre-S2 (Budkowska et al., 1993) have been proposed as possible receptors for HBsAg, a finding which questioned the role

Table 1.1 Proposed HBV-binding receptors.

Binding Protein	Function	Mediator
Albumin receptor	Endocytosis	Polymerised albumin
Transferrin receptor	Endocytosis	Not known
IgA receptor	Endocytosis	IgA
LDL receptor	Endocytosis	Apolipoprotein H
Asialoglycoprotein receptor	Endocytosis	Not known
Interleukin-6 receptor	Signal transduction	Not known
Annexin V	Cell adhesion	Direct
GAPD	Metabolism	Direct

of IgA receptors as an HBV binding site, previously proposed by Pontisso et al., 1992. Apolipoprotein H, found on lipoproteins and taken into hepatocytes after binding to small HBsAg, has been also suggested (Mehdi et al., 1994; 1996). Furthermore, in an attempt to identify cell surface receptor(s) for hepadnaviruses, several proteins which have binding sites for the HBV pre-S region have been proposed, including glycoproteins of 170-180 kd (Kuroki et al., 1994; Tong et al., 1999), 120 kd (Li et al., 1996), 80 kd (Ryu et al., 2000) and 44 kd (De Falco et al., 2001).

Among studies performed for identification of receptor(s) in the initiation of infection by HBV, the role of Annexin-V has been supported by several observations (De Bruin et al., 1995, 1996). Annexin-V, previously referred to as Endonexin II, is a member of

the family of Ca2+ dependent phospholipid-binding proteins present on human hepatocyte plasma membranes (Neurath et al., 1994; Gong et al., 1999). Annexins have also been found to be involved in the initial step of cytomegalovirus and of influenza virus (Gong et al., 1999) and possibly hepatitis Delta virus attachment (De Bruin et al., 1994). Others have argued an important role for human Annexin-V (hA-V) in the cell nucleus (Mohiti et al., 1997). In addition, species-specific distribution of the HBsAg binding annexin V apparently correlates with the species tropism of HBV (De Bruin et al., 1996).

The receptor-ligand relationship between hA-V and small HBsAg has been studied by using antibodies against hA-V which inhibit the binding of HBsAg to intact human hepatocytes (De Meyer et al., 1997). Spontaneous development of anti-idiotype (and, thus, anti-HBsAg) antibodies in rabbits immunised with hA-V and in chickens immunised with the $F(ab)_2$ fragment of rabbit anti-hA-V IgG, are able to compete with hA-V for the binding to HBsAg (De Bruin et al., 1995). Neurath and colleagues (1994) assumed that hA-V and apolipoprotein H bind sHBsAg via the lipid component of HBsAg and not HBV protein specifically.

1.2B Viral entry and uncoating

Although the exact post-binding steps are still unclear, several mechanisms have been suggested for viral entry. There is evidence from DHBV that entry proceeds by a pH-dependent mechanism (Rigg and Schaller, 1992). Endocytosis has also been described (Offensperger et al., 1991; de Bruin et al., 1995). Additionally, Lu et al. (1996) showed that HepG2 cells internalized HBV after proteolytic cleavage of the preS domain by V8 protease. This exposed a fusion domain within the envelope protein of HBV that enabled the viral fusion with host cell membranes. It is also not clear in what form the viral genome is transported into the cell nucleus since the diameter of core particles is at the maximal limit for transport through the nuclear pore (this issue

will be discussed in detail in section 1.3.1B). The HBV genome, once in the nucleus is repaired to form its covalently closed circular (ccc) DNA conformation. This process requires the following modifications: the removal of its terminal structures (from both strands); repair of the single-stranded gap region (by completion of the plus-strand); and covalent ligation of the DNA termini. Host cell enzymes are most likely responsible for all of these reactions (Köck and Schlicht, 1993).

1.2C Transcription and transcription regulation

Viral ccc HBV DNA is transcribed by host RNA polymerase II. It creates both an RNA template for HBV replication and leads to the synthesis of all viral mRNA transcripts (Summers and Mason, 1982). There are at least four groups of mRNA species produced in hepatocytes (reviewed by Seeger and Mason, 2000) (Fig 1.1). These are exported into the cytoplasm where translation of viral proteins, assembly of viral particles and genome replication occurs. The largest, 3.5-kb, which is greater than genome length, and is termed the pregenome RNA (pgRNA), covers the entire genome and is terminally redundant. It is reverse transcribed to DNA. It also acts as a message for HBcAg and polymerase protein synthesis. The second, the pre-core mRNA, is used to translate HBeAg. Pre-core mRNA consists of longer and shorter mRNAs, with start sites separated by~18 nucleotides (Fig 1.2). For HBeAg synthesis and secretion, pre-core mRNA needs to be transcribed followed by translation of the pre-core and core regions (reviewed by Okamoto et al., 1997). The third set of mRNAs, produce the large, middle, and small S proteins. They are initiated at different initiation sites. It is unclear if the fourth (0.7kb) makes HBx protein; it may be that the pregenome message is also used in this capacity.

It is known that each gene of HBV has one or more promoters regulating its activity and that these promoters are in turn regulated by one or both of the viral enhancer elements, Enh1 (nt 970-1240) and Enh2 (1627-1774), that are located upstream of the

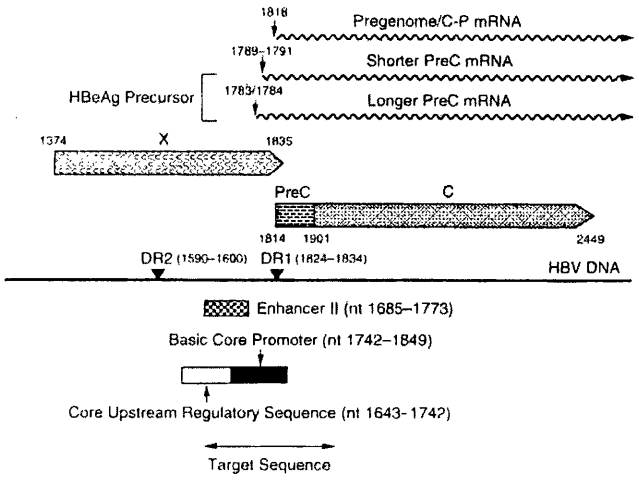


Fig 1.2. Pregenome and pre-C mRNAs of HBV. Start positions of mRNAs are indicated in relation to various genes and elements regulating transcription. Positions of DR1 and DR2 are indicated. Taken from Okamoto et al, 1997.

core promoter (Fig 1.2) (reviewed by Seeger and Mason, 2000). To date, four promoters (S1promoter (S1P), S2 promoter (S2P), C promoter (CP), and X promoter (XP)) have been identified to be responsible for the transcription of the HBV mRNAs (Siddiqui et al., 1986; Treinin and Laub 1987; Honigwachs et al., 1989). Of these four promoters, S2P (nt 3045-3180) is the least liver specific. On the other hand, S1P (nt 2710-2800) and CP are highly liver specific and cells in which CP is inactive can't support HBV DNA replication, as the production of pgRNA which is the template for reverse transcription is mostly under the control of this promoter (Seeger et al., 1989). XP is partly overlapped with Enh1. In addition to the cellular factors that are important for X gene transcription, HBx was shown to have a role in enhancing XP activity by binding a 20bp element at its 5' end (Takada et al., 1996).

A problem of particular importance with a circular DNA template is termination of transcription, especially for the pregenome and precore mRNAs. For these, transcription must pass through the polyadenylation signal on the first pass in order to make a full-size, terminally redundant RNA. Studies with DHBV indicated that this

termination of transcription is facilitated by a sequence on cccDNA, named PET (positive effector of transcription), located between the transcription initiation site and the poly (A) signal (Huang and Summers, 1994). PET somehow facilitates passage of transcription complexes through the termination site on the first but not the second passage. A second signal sequence, NET (negative effector of transcription), which is needed for termination, is suppressed by PET. If NET is deleted, longer transcripts accumulate due to additional transits around the cccDNA template (Beckel-Mitchener and Summers, 1997).

Apart from transcription regulation, regulation of viral gene expression also occurs at the level of translation. The pregenome serves as the mRNA not only for the viral core protein but also for the viral polymerase, which initiates from an AUG located in the distal portion of the core gene, although not in the same open reading frame as core. However, since core particles are assembled from 240 core protein subunits (Crowther et al., 1994), pgRNA may serve as an mRNA for translation, on average, of approximately 200 to 300 core polypeptides before allowing the translation of a polymerase polypeptide. Since the polymerase preferentially binds to the 5' end of its own mRNA to initiate reverse transcription and packaging (Bartenschlager et al., 1990; Wang and Seeger, 1993), synthesis of the polymerase is sufficient to stop further translation of the pregenome.

Spliced transcripts have also sometimes been detected, but there is so far no evidence that RNA splicing plays an important role in HBV gene expression.

1.2D DNA replication

Viral DNA replication is accomplished by reverse transcription of an RNA intermediate, after encapsidation of RNA template. Of many viral transcripts produced in infected cells, only pregenomic RNA is encapsidated. Encapsidation of RNA template with the viral polymerase into the core is followed by sequential

synthesis of two viral DNA strands. Minus-strand is made first from the RNA template, accompanied by degradation of the latter, followed by the plus-strand synthesis using the newly synthesized minus-strand as a template.

1.2D-I RNA encapsidation

The sequences of pgRNA required for encapsidation have now been defined. For HBV, the 5' 100 nucleotides suffice to confer efficient encapsidation upon heterologous transcripts; this region is termed ε. As shown in figure 1.3, this region contains several inverted repeats that can be folded into a stem loop structure that appears to be phylogenetically conserved (review by Ganem et al., 1994). Extensive mutational studies have suggested that this region of the structure is functionally important for encapsidation (see section 1.7.3B), as in the loop region at the top of the structure, even minimal base substitutions are strongly deleterious to encapsidation (Pollack and Ganem, 1993).

1.2D-II DNA synthesis

DNA replication starts with the binding of the viral polymerase to the ε bulge at the 5' end of pgRNA (Figure 1.4). The polymerase then primes DNA synthesis, using a tyrosine in its own aminoterminal protein domain (TP) as a primer and ε as a template. After three or four nucleotides the DNA synthesis is arrested. Further DNA synthesis requires three switches; one during minus-strand synthesis and two during plus-strand synthesis. The first template switch occurs shortly after the priming reaction. The primed 5' ⟶ 3' pgRNA-Pol complex is translocated to a complementary sequence (UUAC) near the 3' end of the pgRNA (Fig 1.4B) where the synthesis of minus-strand DNA resumes. The minus-strand is then extended to the 5' end of the pgRNA (Fig 1.4C and D). An active role for the polymerase in minus-strand DNA transfer has been recently described in DHBV as deletion of aa 79-88 in

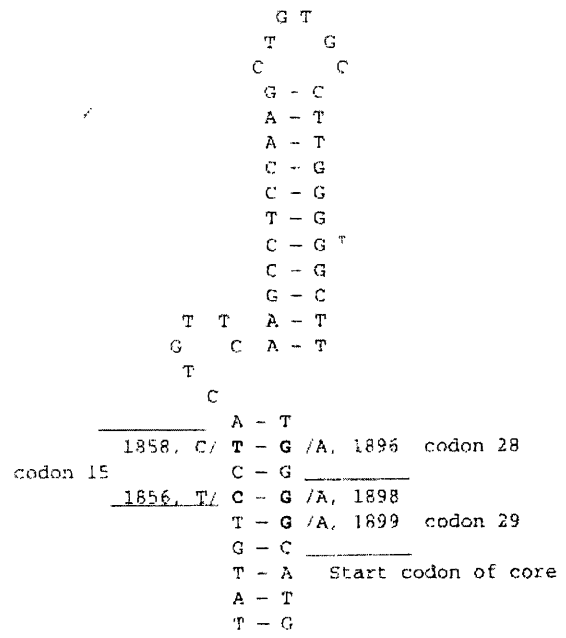


Fig 1.3. Predicted stem loop structure of pregenome encapsidation sequence (ε) and location of the common pre-C HBV mutants. Taken from Lok et al, 1997.

the terminal protein domain specifically inhibited the minus-strand transfer reaction(Gong et al, 2000). Accompanying or shortly following the minus-strand synthesis, the pgRNA template is degraded by the RnaseH activity of Pol with the exception of a short terminal oligoribonucleotide (Fig 1.4D). This oligoribonucleotide is then translocated, in the second switch, to DR2 near the 5' end of minus-strand DNA where it serves as the primer for plus-strand DNA synthesis (Lien et al., 1986; Loeb et al., 1991). Following this second transfer, the plus-strand DNA synthesis continues to the 5' end of the minus-strand DNA (Fig 1.4H). However, in approximately 10% of cases the second switch from DR1 to DR2 does not occur, leading to an in situ priming reaction (Staprans et al., 1991) (Fig 1.4E). Thus a double strand linear DNA genome rather than relaxed circular DNA (rcDNA) is produced. A third template switch occurs once the plus-strand DNA synthesis reaches the 5' end of minus-strand DNA. The minus-strand DNA template contains a 7-9 nt terminal redundancy (r). This redundancy is required for the third template switch in which the 3' end of the nascent plus-strand is translocated to the 3' end of minus strand (Loeb et

al., 1997) (Fig 1.4I) to circularize the genome and permit resumption of the plus-strand synthesis. The synthesis of plus-strand is only partially completed resulting in a noncovalently closed, partially double stranded, circular DNA genome (rcDNA).

1.2E Viral assembly

Nucleocapsids containing rcDNA, can be then either enveloped at an internal cellular membrane and be actively secreted as mature virions or transported back into the nucleus, where the genome is repaired to yield cccDNA, to amplify the intracellular genome pool (Tuttleman et al., 1986). Thus, cccDNA amplification is caused not only by a reinfection of the cells but by re-entry of the viral genomes (similar to artificial transfection) from the mature core particles (Seifer et al., 1993, 1998; reviewed by Nassal, 1999) which assemble in the cytosol. Persistence of the virus in infected cells during chronic HBV may therefore be dependent on this intracellular conversion pathway which generates, maintains, and regulates the pool size of transcriptionally active cccDNA molecules (Tuttleman et al., 1986; Serinoz et al., 2003). It is likely that these nuclear levels of cccDNA are specifically regulated to maintain amounts that are sufficient for virus production but are not toxic to the cells. This amplification could be regulated by either a cell-directed mechanism or virus-directed negative feedback inhibition (Fig 1.5). Negative control over the conversion of rcDNA would be effected through production of viral gene products at both transcription and translation levels, whose concentration in the cell is dependent on the size of the cccDNA pool (reviewed by Seeger and Mason, 2000). These gene products would then block further synthesis of cccDNA directly, or would direct assembly of cytoplasmic nucleocapsid exclusively into mature virus (Tuttleman et al., 1986). Other authors showed that high amounts of L protein in the ER lumen blocked viral maturation (Bock et al., 1999).

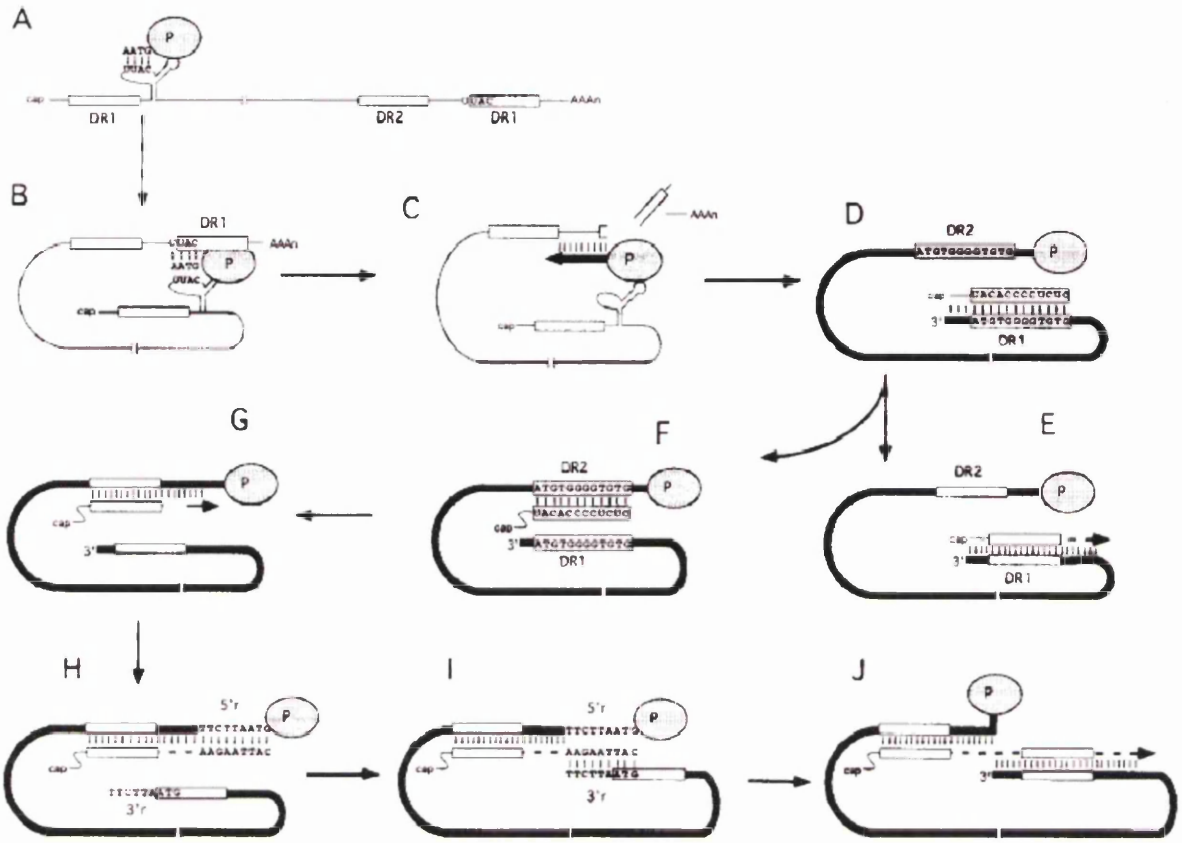


Fig 1.4. HBV genome replication.

The thin line stands for pgRNA, dashed line for the plus strand and the bold line for the minus strand. The direct repeats are presented as boxes labeled DR1 and DR2. The UUAC involved in the minus strand template switch are shown. The dashed circle represents the viral P protein, which binds the ε bulge at the 5' end of pgRNA. The sequences of the terminal redundancy of minus strand DNA are shown as 5'r and 3'r. Basepairing is shown as hatch marks. B, Minus strand template switch. C, Elongation of the minus strand. D, completion of minus strand synthesis and generation of the plus strand primer. E, generation of duplex linear genome. F, plus strand primer translocation. G, initiation of plus strand synthesis. H, plus strand synthesis continues to the 5' terminus of the minus strand DNA. I, circularization template switch. J, generation of a relaxed circular DNA genome. Taken from Harvet and Loeb, 1997.

1.2E-I Surface-gene Regulation

L protein, in addition to S protein, is also required for the assembly and production of mature virions; however, most of the preS region is dispensable for this process. The L polypeptide when synthesised in an in vitro system, bound firmly to HBcAg, indicating that interaction between HBcAg and the pre-S region of the L polypeptide is important for virus morphogenesis (Dyson et al., 1995). Using electron microscopy, Kamimura et al, (1981) demonstrated that in the presence of L and S proteins, cores could bud into the intermediate compartment between the endoplasmic reticulum and the Golgi complex, and exit the cell as enveloped virions through the secretary pathway.

It has long been known that maturation of HBV virions requires interaction between HBV S and core proteins on the ER. This interaction also appears to regulate the nuclear transport of HBcAg: Yeh et al, (1994) described that nuclear transport was inhibited to an undetectable level in COS-7 cells when S proteins were co-expressed with core protein. Further analysis revealed that L but not S protein reduced the amount of the core protein in the nucleus and was crucial for this inhibitory effect. This effect was attributed to the amino-terminal sequence of L which interacted with core protein (Yeh et al., 1994) and specifically to aa positions between 19-24 and 63-65 aa in the pre-S domain which are essential in mediating interaction with the nucleocapsid (Dyson et al., 1995).

There is an assumption that the pre-S proteins regulate viral cccDNA levels via negative feedback (Summers et al., 1990). It is known that overexpression of the L protein can block secretion of subviral particles as well as virions (Hadziyannis et al., 1987; Bock et al., 1999). These authors therefore suggested that early in infection, when there are low amounts of surface proteins, mature nucleocapsids are shunted

into the nucleus and the newly synthesised viral genome converted into cccDNA. The number of cccDNA molecules increases as infection progresses, leading to an increase of viral RNA synthesis. The viral envelope proteins are subsequently produced at high levels and exert a negative feedback on cccDNA amplification by redirecting the nucleocapsids into the virion formation pathway (Bartholomeusz and Locarnini, 2001). Failure of regulation of the cccDNA copy number by the envelope proteins results in hepatocyte death (Lenhoff et al., 1994). On the other hand, in the late stage of infection, when the L protein levels increase, export of core particles (as the mature virions) occurs (Yeh et al., 1994).

1.2E-II Cell cycle Regulation

The effect of cell cycle regulation on intracellular core protein has been studied extensively by Yeh and co-workers. They showed in transfected cell lines as well as human hepatocytes that the core protein localization was largely in the nucleus in the G1/0 phase but in the cytoplasm in the S phase (Yeh et al., 1993, 1995, and 1998). They also found that nuclear localization of HBV DNA (RC form) occurred only in the G1 phase and that prolonged aphidicolin treatment was required for cccDNA formation. Further, they showed that when cells were grown to confluence and arrested in the G0/G1 phase, the nuclear transport of core protein was significantly increased. These results revealed that hepatocytes in different cell cycle phases are associated with different directions of viral DNA traffic. In the G1 phase, the nuclear

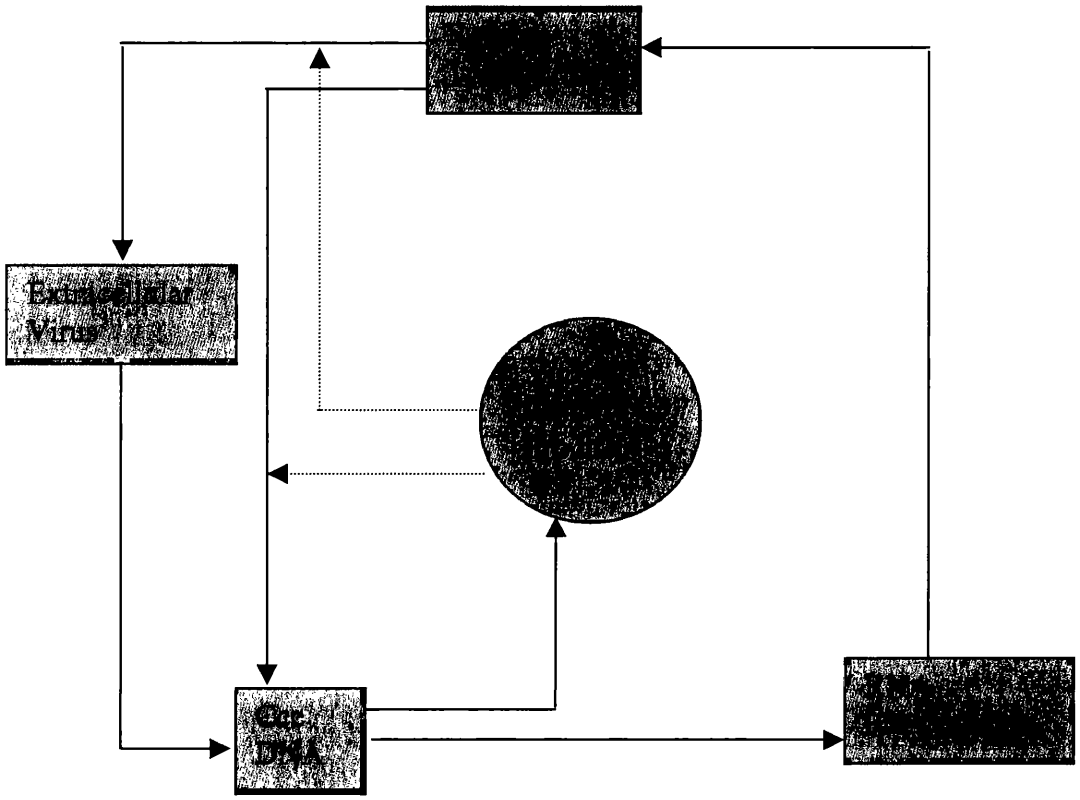


Figure 1.5: cccDNA pool formation during the early stages of HBV infection. Dotted lines indicate the inhibitory effects of viral gene products on cccDNA formation. Taken from Tuttleman 1986.

RC viral DNA is processed (excision of polymerase, ligation, DNA repair, etc) to cccDNA. HBV DNA is recycled into the nucleus to form cccDNA, so only a small amount is destined for viral maturation. This small but constant out-flow of virions will eventually result in a high concentration of serum HBV DNA. The accumulation of cccDNA clearly also results in an increase of total intracellular HBV DNA, which can be used for virion maturation. In contrast, in the S phase, nuclear entry of HBV DNA does not occur, so the majority of the viral DNA is destined for virion maturation. Of note is that a prolonged G1 phase is not the only requirement for cccDNA formation, and other cellular factors (which are species specific) must be involved (Yeh et al., 1998). This assumption was based on the observation that, in transgenic mice, no cccDNA could be found in nuclei, although most hepatocytes were in the G1/0 phase (Guidotti et al., 1994a). Conversely, a constantly regenerating liver (as seen in aggressive hepatitis) is unfavorable for HBV replication because the G1 phase may not be long enough for ccc DNA accumulation. This explained a recent observation that HBV DNA replication is enhanced in quiescent hepatocytes (Chu et al., 1997). On the other hand, Guidotti et al (1994, 1997) demonstrated that in the absence of certain cytokines, HBV can efficiently replicate in dividing hepatocytes, suggesting that HBV gene expression is relatively resistant to signals that upregulate or downregulate the expression of cellular genes during liver cell regeneration. They also claimed that the nuclear membrane allows translocation of only the non-particulate core protein. Further, in separate studies carried out by Yeh et al (1990, 1993) and Eckhardt et al (1991), it was suggested that this translocation which was dependent on NLS (nuclear localization signal) at the C-terminus (as will be discussed next), was regulated through differential phosphorylation of the serine on the SPRRR motifs at different cell cycle stages (see below).

1.3 HBV viral proteins

1.3.1 C protein

1.3.1A Structure

The regular display of multiple identical structural features on the HBV core particle surface is crucial for its exceptional immunogenicity and ability to act as a T cell independent antigen (Nassal et al, 1992a). Two classes of the core particle have been identified: 120 and 90 dimeric subunits in T =4 and T =3 arrangements.

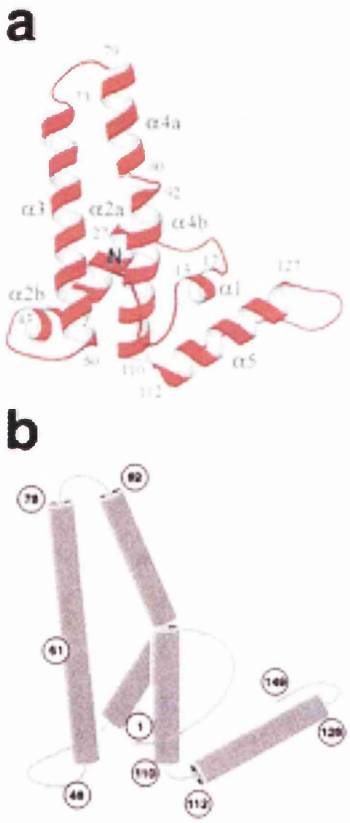


Figure 1.6. The Tri-dimensional model of the HBV capsid protein dimer on the basis of crystal structure (a) and from electron cryomicroscopy (b). Central α-helices (boxes) and the loop covering aa residues 78-83 are shown. C and N indicate C-Terminal and N-Terminal domains, respectively. Taken from Wynne et al, 1999.

The building block of the capsid structure is a T-shaped dimer (inverted), the stem of which interfaces with the other dimer and protrudes outwards as a spike; the cross-

pieces pack together to form the contiguous shell (Conway et al., 1998a) (Figure 1.6). This characteristic folding of core protein has been defined by its crystal structure (Wynne et al., 1999) (Fig 1.6a) and by cryoelectron microscopy (Böttcher et al., 1997; Conway et al., 1998b) (Fig 1.6, b). The two dimers run from the inside of the capsid to the tip of the spike (Kratz et al., 1999), which consists of a four α-helix bundle. Amino acid residues 78-82 form a non-helical loop connection between two alpha-helical segments (Bringas et al., 1997) at the tip of the capsid spike (Figure 1.6b). The region 128-149 forms an arm-like structure that plays a critical role in capsid assembly (Hui et al., 1999). This arm or loop makes a small protrusion on the surface of the shell and provides the next most exposed and accessible feature on the surface. It is known that the principal antigenic sites of the protein correspond to these residues. The interaction between dimers to form the icosahedral shell is mediated by the C-terminus. This region is positioned within the shell interacting with nucleic acid (Bottcher et al., 1997).

With a 30 or 34-nm diameter, the core encapsidates the circular DNA genome. HBc protein packages the pgRNA and the viral polymerase. Early studies indicated that the C-gene encoded two polypeptides. Initiation of translation at the first AUG of the pre-C region in a "pre-C-mRNA" produces a 25-kd pre-core/core precursor protein (p25e), which after removal of the leader sequence at the amino-terminus (first 19 residues) and the 34 C-terminal aa (Takahashi et al., 1983), is secreted as a soluble 17-kd protein (p17e), HBeAg (Fig 1.7). The C-terminus marks the processing site in the sequence (146) T V V R R R G (153) (Figure 1.8) that is selectively cleaved during HBeAg secretion in vivo, and the 10 amino acids remaining from the pre-C sequence are essential for the correct folding of HBeAg (Nassal et al., 1993). Non-carboxy-terminally processed pre-core/core gene products comprised of 193 aa from

position –10 (the site of signal cleavage during HBeAg secretion) to 183, are designated p22e (Figure 1.7) (Salfeld et al., 1989).

Initiation at the second AUG in the pgRNA, leads to the synthesis of HBcAg, a 183 aa, 21 kd, protein (p21c). Icosahedral shells of two sizes are produced, containing 180 (T=3, where T is the triangulation number and the shell contains 60T subunits), or 240 (T=4) subunits, which assemble to form the viral core (Wingfield et al., 1995). Core protein is 183 or 185 residues long depending on the subtype. Despite p17e and p21c sharing a large stretch of identical primary sequence, HBeAg and HBcAg are serologically distinct (see section 1.6). HBeAg is always present in a non-particulate state in the serum of HBV infected individuals, whereas HBcAg occurs only in the enveloped virus or inside HBV infected hepatocytes (Salfeld et al., 1989).

Immunofluorescence studies strongly suggest that expression of the pre-C gene determines the attachment of the nascent p25e to the membrane of the ER. The affinity of the pre-C protein to cellular membrane and its amino terminal sequence resembles the properties of signal peptides known to transport proteins through cellular membranes, an event which is a pre-condition for acquiring the viral envelope and subsequent secretion (Ou et al., 1986; Uy et al., 1986; Ballare et al., 1989; reviewed by Okamoto, 1997). This translocation requires the first 19 residues of the pre-core protein as a signal peptide to direct it to the ER (Garcia et al., 1988), where they are cleaved with the formation of a pre-C protein-derivative, p22e (Ou et al., 1986; Garcia et al., 1988). P22e is then released in the ER and Golgi apparatus with further hydrolysis at the arginine-rich C-terminus prior to secretion (p17e); some p22e, however, is released into the cytoplasm and translocates to the nucleus (Figure 1.7). For this, it requires the first 10 aa residues as a nuclear localisation signal (Garcia et al., 1988; Ou et al., 1989; Yeh et al., 1990).). Thus, the pre-C sequence

alters profoundly the properties of the viral core protein (Uy et al., 1986); if the proximal pre-C sequence is replaced by a random sequence, the resulting mutant protein cannot be transported into the nucleus (Yeh et al., 1990). Thus, the 29 aa residues preceding the core protein might function as a signal peptide for secretion of HBeAg or as a nuclear transportation signal for p22e. (Yeh et al., 1991; reviewed by Okamoto, 1997). Both P22e and core protein (p21c) contain nuclear localisation sequences, but in different positions (for core protein nuclear localisation signal see section1.3.1B); there are 10 additional residues at the amino-terminus of P22e and therefore, the cellular regulation of nuclear transport of these two proteins may be different (Yeh et al., 1996). Therefore, HBcAg expressed without the pre-C sequence (p21c) accumulates as a high molecular weight complex in the cytoplasm; in contrast, the HBcAg produced with the pre-C sequence can be a nuclear protein. Thus, the fact that a significant proportion of p22e is retained within the cell nucleus indicates that p22e plays an important role in the HBV life cycle (Ou et al., 1989; Aiba et al., 1997).

1.3.1B C protein C-terminal domain and its phosphorylation

The core protein consists of two structurally and functionally separate domains: the sequence from the N-terminus to about aa position 149 is by itself sufficient for assembly, known as the "assembly" domain, and a "protamine" domain at the C-terminus, which is involved in pregenome encapsidation and modulates the activity of polymerase for reverse transcription of the RNA pregenome and binding to the elongating DNA (Hui et al., 1999). If the C-terminus is truncated (after aa 140), the core protein still assembles into a shell, but no longer packages RNA (Bottcher et al., 1997).

The C-terminus 33 aa of core contains 16 (adr and ayw subtypes) or 17 (adw subtype) arginine residues; 14 of them are clustered into 4 arginine repeats, in line with its role

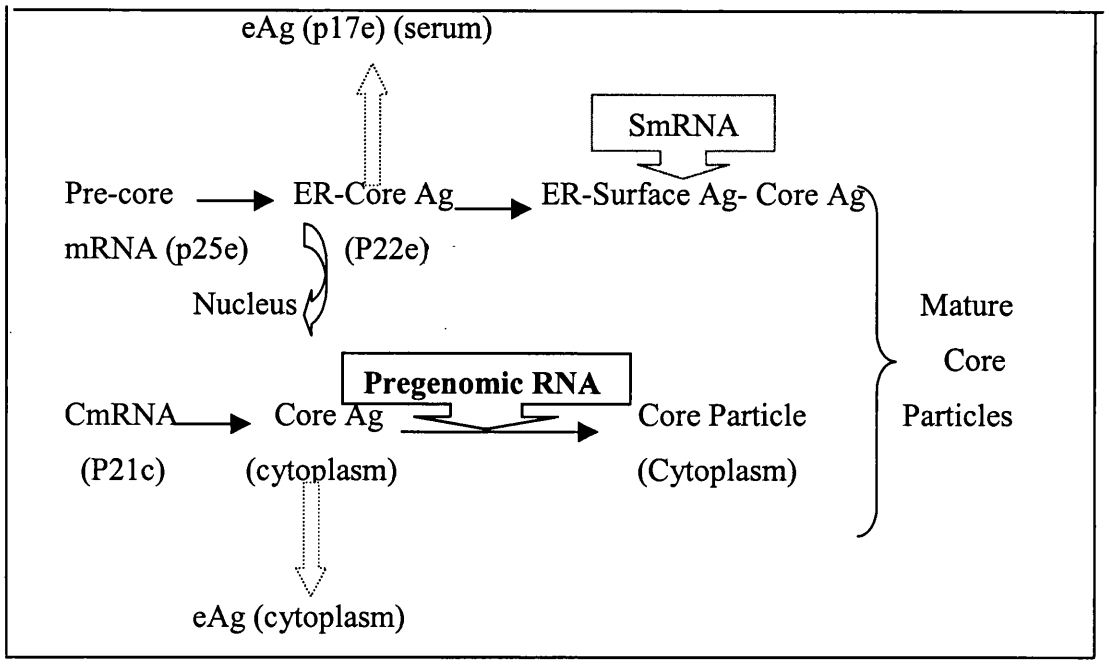


Figure 1.7: Hypothetical model for the assembly of HBV particles.

P22e processed from p25e binds to the ER (membrane-bound core protein). There are alternatives: molecules may be transported into the lumen, perhaps processed, and secreted as HBeAg (p17e). HBcAg translated from core-mRNA (p21c) remains in the cytoplasm and interacts with pregenomic RNA to form core particles. Some of the p21 exists in the HBeAg form. The core particles containing HBV RNA or DNA may then interact with HBcAg and HBsAg and become encapsidated, presumably via a budding process, with consequent formation of mature viral particles. CmRNA and SmRNA, represent core mRNA and surface mRNA, respectively.

in nucleic acid binding (Figure 1.8) (Nassal et al., 1992b). These regions of consecutive basic aa are similar to nuclear localisation sequences identified in several nuclear proteins: (Eckhardt et al., 1991); influenza viruses (Greenspan et al., 1988); Ela in adenoviruses (Lyon et al., 1987); yeast histone 23 protein (Moreland et al., 1987); yeast ribosomal protein L29 (Underwood et al., 1990); large-T protein of polyoma virus (Richardson et al., 1986); and SV40 large-T Ag (Kalderon et al., 1984).

The C-terminal domain seems crucial for HBeAg production. HepG2 cells infected with a recombinant vaccinia virus expressing a pre-C protein truncated by 34 aa at the HBeAg C-terminus secreted about 20 times less HBeAg than those infected with a recombinant expressing the entire precore protein (Schlicht et al., 1991). Thus, low secretion efficiencies of C-terminally truncated precore proteins are due to a slow down in intracellular transport of the HBeAg precursor (Carlier et al., 1995). Taking all the data together, one possible function of the pre-C protein might be to transport progeny viral genomes from the cytoplasm into the nucleus, where they can direct the synthesis of additional RNA and eventually increase the number of progeny particles produced by the infected cells. The relatively higher expression of pre-C over core results in translation of a greater amount of pre-C protein. This is then either secreted as HBeAg or transported to the nucleus as nuclear HBcAg (Ou et al., 1990). This hypothesis is in keeping with the clinical observation of high levels of serum HBV DNA/HBeAg and the predominant localization of HBcAg in the nucleus during this phase (see section 1.7.3D) (Hadziyannis et al., 1983; Wu et al., 1987; Hsu et al., 1987; Akiba et al., 1987; Chu et al., 1997). In contrast, expression of the pre-C protein is relatively lower than that of core in the low replicative phase.

HBcAg is a nuclear, cytoplasmic, and membrane-associated phosphoprotein (Roossink et al, 1987) that shuttles between cytoplasm and nucleus during the different stages of chronic infection (Chu et al., 1987; Hsu et al., 1987; Akiba et al., 1987). It is generally believed that viral assembly of hepadnaviruses occurs in the cytoplasm as the capsid proteins are localised there (Yeh et al., 1996). Since diffusion of molecules smaller than 60 kd is allowed through a 9-11 nm aqueous channel, transport of large proteins requires one or more of the signal sequences to induce opening of the nuclear pore complex up to a diameter of about 26 nm (Pante et al.,

1993). Thus, the 30-34 nm HBV core particle is excluded (reviewed by Kann, 1997). Also, the nuclear membrane itself is a natural barrier for intact core particles in transgenic mice, and core particles can only migrate to the cytoplasm during mitosis when the nuclear membrane is disintegrated (Guidotti et al., 1994a). This allows core protein, which had formed capsid-like particles in the nucleus, to be released back into the cytoplasm (Liao et al., 1995). After cytoplasmic disassembly of core particles, HBcAg remains associated with the viral genome. This nucleoprotein complex is then targeted to the nuclear pore complex (NPC) by the use of the core protein nuclear localisation signals, where it is translocated through the pore complex into the nucleus (Bock et al., 1996).

During the early phase of infection, because of the quiescent state of hepatocytes, nuclear entry of the core will be facilitated. This nuclear localisation may be important for transporting the HBV DNA into the nucleus for initiating viral mRNA transcription. Also, the nuclear localisation of the core is important for transporting the viral DNA into the nucleus for cccDNA amplification. Late during infection, production of large viral envelope protein inhibits the former, redirecting viral nucleocapsids into enveloped virus particles, which are exported from the cell (Summers et al., 1990). Thus, control of cccDNA amplification is necessary for the ability of hepadnaviruses to infect cells persistently without cytopathic effects.

Within the last few years, considerable progress has been made in understanding the role of the C-terminus arginine-rich clusters of the HBV core protein. HBcAg can be phosphorylated in vivo (Roossinck et al., 1987) and in vitro (Albin and Robinson, 1980; Gerlich et al., 1982; Petit et al., 1985). While the core protein phosphorylation may be important for pre-genomic RNA packaging, its dephosphorylation may be important for viral DNA replication (Lan et al., 1999). It may even regulate

maturation of the virion (Pugh et al., 1989), as Gazina et al, (2000) reported that HBV core protein was phosphorylated prior to assembly into nucleocapsids. Previous studies have shown that core protein contains one or more NLS (Yeh et al., 1990; Eckhardt et al., 1991) which may not be exposed on the surface of the core particles (Kann et al., 1999). However, Eckhardt et al, (1991) proposed two nuclear localisation sequences (Fig 1.8), suggesting that at least a portion of each of these regions might be present on the surface of the protein. Collectively, authors presumed that HBcAg can be recognised by the NLS receptor(s) and subsequently transported to the nucleus by exposure of NLS(s) (Eckhardt et al., 1991; Kann et al., 1999).

Numerous studies have shown that core particles, purified from infected liver or circulating virions, display protein kinase (PK) activity (Albin and Robinson, 1980; Gerlich et al., 1982; Petit et al., 1985). Some suggest a host origin of this PK (Lanford et al., 1990). This is supported by the lack of a protein kinase consensus sequence in the viral genome (Hanks et al., 1988; Schlicht et al., 1989). Studies on E.coli-derived core particles have shown that phosphorylation of the core subunits was necessary for exposure of the nuclear binding signal (NLS), and inhibition by specific peptides indicates that the COOH-terminus was directly involved (Eckhardt et al., 1991; Liao et al., 1995; Kann et al., 1999) through a direct interaction with a kinase or cellular factor (Hui et al., 1999).

The COOH-terminal domain of the core protein which is highly positively charged overlaps and interferes with the nucleic acid binding domain (Fig 1.8) (Petit et al., 1985). In unphosphorylated cores, cryoelectronmicroscopy has shown that this domain is located inside the cores close to small holes in the capsid wall (Zlotnick et al., 1997); it binds to nucleic acid (Petit et al., 1985; Nassal et al., 1992b; Liao et al., 1995) and also provides the phosphorylation site(s) for the protein kinase trapped in

the central cavity (Machida et al., 1991; Liao et al., 1995; Kann et al., 1997). Removal of the last 36 C-terminal aa completely abolished phosphorylation and genome replication but did not interfere with particle formation or genome packaging (Schlicht et al., 1989). Since the phosphorylation sites on core proteins interfere with the encapsidation of RNA (Kann et al., 1995), this process is most likely linked with DNA synthesis. Thus, core protein phosphorylation may be an essential step during the generation of a replication-competent Hepadnavirus capsid (reviewed by Cooper, 2003).

The three proposed phosphorylation sites in the HBV core protein sequence have been mapped exclusively to the 7 serine residues in the arginine-rich C-terminus of the core and pre-core proteins, six of them in a highly conserved region (Figure 1.8) (Gerlich et al., 1982; Machida et al., 1989; Yeh et al., 1990; Liao et al., 1995). These serine residues reside in the three adjacent SPRRR repeats and overlap the nuclear localization signal (Figure 1.8). Sequences of other nuclear localisation signals in other viruses such as WHV, GSHV (Yeh et al., 1990), and DHBV (Walker and Lipkin, 2002) are also rich in arginine or lysine residues or both. Certain serine residues at amino acid positions either 155, 162, 168 and 170 (Argos and Fuller, 1988; Machida et al., 1991) or 145-156 (Eckhardt et al., 1991) were proposed as candidates for phosphorylation, being the consensus sites for different kinases (cdc2, cdk2, PKC, SRPR1 and SRPR2) (Kann et al., 1994; Liao et al., 1995; Daub et al., 2002). The requirement for the phospho-acceptor sites to be juxtaposed next to proline has been conserved during evolution.

Mutations of three phosphorylation sites (S157, S164, and S172) to alanine (to mimic dephosphorylation) abolished pre-genomic RNA packaging. Conversely, the mutation of these three phosphorylation sites to glutamic acid (to mimic phosphorylation), had

only a slight effect. However, such mutations were not able to support viral DNA replication (Lan et al., 1999). In DHBV alanine could functionally replace serine at the sites of replacement, but not for the phosphoreceptor function. This defect was corrected by aspartic acid substitution (the assumption is that aspartic acid mimics phosphoserine rather than serine) (Yu et al., 1994). The same results were obtained for S176 and S178 after substitution by alanine (Daub et al., 2000). Taken together, it seems that phosphorylation of core subunits induces a conformational change. As a consequence, the COOH terminus is then exposed on the particle surface and serves as a nuclear pore targeting signal (Kann et al., 1999).

In contrast, observations that the phospho-core protein was detected in the cytoplasm but not in the nucleus, suggested that phosphorylation of the core protein has a negative effect on nuclear localisation and that the cell cycle regulated nuclear localisation. The core protein accumulates primarily in the nucleus when entry of the cell cycle into S phase is blocked by aphidicolin; removal of the chemical resulted in the entry of the cells into S phase and the accumulation of the core protein in the cytoplasm. It is assumed that lack of nuclear core protein during the S phase may be the result of activation of a cellular kinase at this particular stage of the cell cycle. Authors found that this kinase could phosphorylate the core protein in vitro. They also suggested the possibility that this kinase and/or its related kinases may negatively regulate the nuclear localisation of the core protein during the S phase (Yeh et al., 1998). Although the data presented here on phosphorylation of recombinant, non-assembled HBV core protein cannot be necessarily transferred to the genuine complete core particle (containing genome, polymerase and kinase), further support is lent by similar observations with the DHBV core particle.

1.3.2 S protein

The principal viral surface antigen (HBsAg) is a protein of 24 kd in virions and subviral particles. There are 100 polypeptide monomers per subviral particle, and in addition host-derived lipid that comprises approximately 25% of the particle mass (Ganem et al., 1987). ORF-S is embedded within the ORF-P (Figure 1.1). It encodes the three envelope proteins which are translated by three in frame AUG codons. They all share their C-terminus sequence and "a" antigenic determinants (see below) (Argos and Fuller, 1988). The three-conserved AUG codons are for large (L; containing pre-S1, pre-S2, and S), middle (M; containing pre-S2, and S), and small (S; HBsAg) proteins. LHBs covers the entire open reading frame, MHBs initiates at an internal site, and the 226 aa sequence of SHBs starts further downstream. All three proteins therefore have the S sequence in common. The pre-S proteins represent a minor component of the circulating antigenic pool (generally less than 1-10% of the total). All three proteins are inserted into the outer envelope of the virion (Argos and Fuller, 1988). These proteins are usually detected serologically as HBsAg (Ganem et al., 1996), and they are involved in receptor binding, viral assembly and secretion. They are also important targets for immune mediated virus elimination (reviewed by Wallace et al., 1997). The S proteins in secreted viral and subviral particles are extensively cross-linked by disulphide bridges; such cross-linking has been shown to occur concomitantly with the assembly and budding of the particle (Huovila et al., 1992). Of 14 cysteine residues presumably involved in this process, four are located in the first hydrophilic region and eight are located in the second hydrophilic region (Mandart et al., 1984; Norder et al., 1992b).

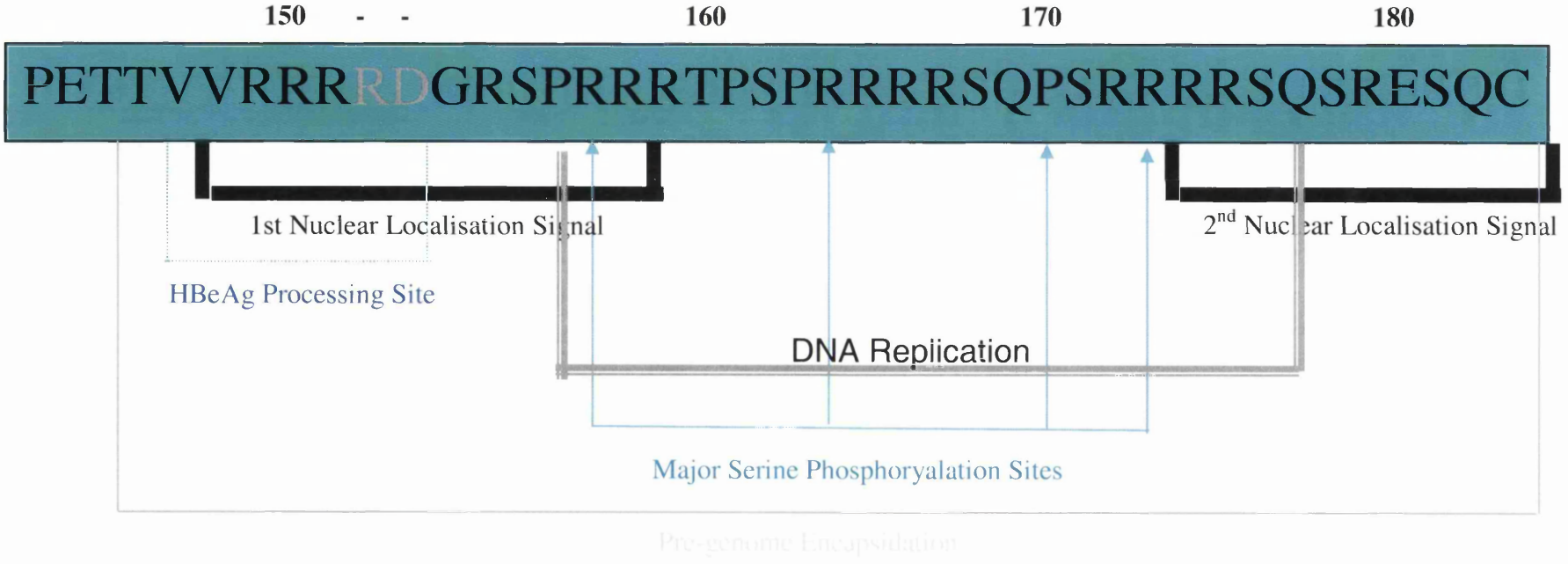


Figure 1.8: A schematic structure of the HBcAg C-Terminus domain representing various functional aspects. Numbers at the top indicate the number of HBcAg C-terminus amino acid residues. RD indicates the two extra amino acids in genotype A, adw2 strains represented in black colour.

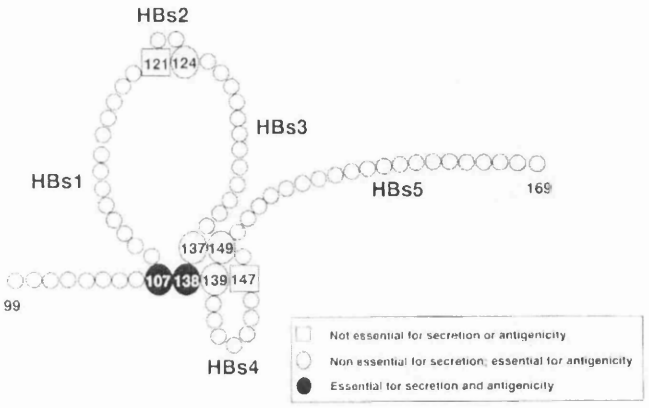


Fig 1.9. Proposed model of the major hydrophilic region (MHR). The five proposed antigenic regions are labled HBs1 through HBs5. Cysteine to cystein disulphide bridges are also shown. Taken from Wallace and Carman, 1997.

The first hydrophilic region, aa 29-79, is on the inner cytoplasmic surface of the viral envelope and thus may facilitate core particle envelopment during maturation of the virion (Prange et al., 1995; Löffler-Mary et al., 2000). Three of the four cysteine residues in this region are conserved among all hepadnaviruses and were shown to be essential for secretion.

The second hydrophilic region (aa 99-147) is exposed on the outer surface of secreted viral and subviral particles. It is termed the major hydrophilic region (MHR) (Fig 1.9) and most of the anti-HBs stimulated by vaccine bind this; substitution of many residues results in loss of immunogenicity to vaccines. This region has a highly complex structure and is very cysteine rich; eight of the 14 cysteine residues in HBsAg are located here and all of them are highly conserved among mammalian hepadnaviruses (Mandart et al., 1984; Norder et al., 1992b, 1993a). Epitopes of the MHR of HBsAg, cluster into five regions (Fig 1.9): HBs1, upstream of aa121; HBs2, between aa121 and 124; HBs3, between aa125 and 137; HBs4, between aa 139 and

147 or 149 and HBs5, from 148 or 150 downstream to 169. Antigenicity of HBsAg is dependent upon this complex structure and variants of this region have been well described (Carman et al., 1990a; Protzer-knolle et al., 1998; Esteban et al., 1999). Moreover, standard HBV subtyping and genotyping is based on the "a" determinant which does not include the entire MHR (see section 1.5)

1.3.3 X Protein

The smallest ORF, the X-gene, encodes a polypeptide of 154 amino acids. HBx is highly conserved among mammal-specific members of Hepadnaviridae (Sohn et al., 2000) and seems to be essential for virus infectivity. HBx has been detected in liver tissue from HBV-infected patients with chronic hepatitis, cirrhotic liver and HCC (Wang et al., 1991; Zhu et al., 1993; Su et al., 1998).

HBx protein activates a wide range of promoters and transcription factor binding sites and multiple functional pathways. It contains the basic core promoter (BCP)/enhancer II complex, which activates transcription of genes, interacts with other transcription factors and is involved in tumorogenic activity (reviewed by Kann et al., 1998) by activation of the P53 tumor suppression protein (Feitelson et al., 1993).

1.3.4 P Protein

The longest ORF encodes the viral polymerase and overlaps the C, S, and X genes (Figure 1.1). It shares homology with the polymerase genes of retroviruses (Ganem et al., 1987). It contains four functional domains which are arranged from the N to the C terminus as follows (Fig 1.10): terminal binding protein (TP), which is covalently linked to the 5' end of the minus-strand of the virion DNA and is necessary for priming minus-strand synthesis; a spacer, which despite being necessary for polymerase function, has no specific function of its own; reverse transcriptase, encoding the nucleic acid-dependent polymearse activity and the reverse transcriptase

activity as it has the conserved YMDD motif.; and RNase H, responsible for digesting the pgRNA. Sequence analysis of the mammalian and avian Polymerase ORF revealed that it is highly conserved among hepadnaviruses (Chen et al., 1992, 1994).

Fig1.10 Functional domain of the polymerase. TP and RT represent terminal binding protein and reverse transcriptase, respectively.

1.4 Transmission and Epidemiology of HBV transmission

HBV is transmitted by a number of routes including transmission from an infected mother to her child (perinatal or vertical transmission), or by percutaneous and mucous membrane exposure to infectious blood and body fluids that contain blood (horizontal transmission). Although hepatitis B surface antigen (HBsAg) has been detected in a wide variety of body fluids, only serum, semen, and saliva have been demonstrated to be infectious. Perinatal and sexual transmission of HBV usually results from mucous membrane exposure to infectious blood or serum-derived body fluids (Alter et al., 1990), and the risk is greatest for infants born to women who are HBeAg-positive, ranging from 70 to 90% at 6 months of age; about 90% of these children remain chronically infected (Xu et al., 1985).

The global prevalence of chronic HBV infection varies widely (Fig 1.11). The prevalence of chronic HBV infection in different areas has been categorized as of high, intermediate or low endemicity (Fig 1.11). Areas of high endemicity (prevalence ≥8%) account for a total of 45% of the global population and include Africa, Asia and the Pacific Basin, the Amazon Basin, the Arctic Rim, the Asian republics previously

part of the Soviet Union, and parts of the Middle East and the Caribbean. Additionally, in parts of Eastern Europe, such as Bulgaria, Romania, Albania and Moldavia, the HBV prevalence is 5-10% of the general population. Areas of intermediate endemicity have a prevalence of chronic HBV infection of 2-7% (Fig 1.11), and account for a total of 43% of the global population. Parts of Southern and Eastern Europe, the Middle East, Japan, parts of Central and South America, are regarded as areas of intermediate endemicity. In most developed parts of the world, the prevalence of chronic HBV infection is less than 2% (Fig 1.11), and include most of Western Europe, Australia and North America that are generally categorized as areas of low endemicity.

The predominant routes of HBV transmission vary according to the prevalence of HBV infection. In regions of high endemicity, transmission from an infected mother to child is the main mode of transmission (Lau et al., 1991; Alter et al., 1996) In regions of intermediate endemicity, individuals of all groups can be infected, although chronic infection is generally caused by transmission during infancy or early childhood. In regions of low HBV prevalence, transmission is primarily horizontal (between individuals). Sexual transmission (either homosexual or heterosexual) in high-risk adults is a main mode of transmission in Europe and North America (reviewed by Alter, 2003), although needle sharing amongst intravenous drug abusers or occupational exposure to contaminated blood and blood products continue to be important (Lee et al., 1997).

1.4.1 Prevention

The primary goal of hepatitis B prevention is reduction of chronic HBV infection and HBV-related chronic liver disease. A secondary goal is the prevention of acute hepatitis B. HBV infection can be prevented by screening blood, plasma, organ tissue

and semen donors. Although such activities can reduce or eliminate the potential risk for HBV transmission, immunization is by far the single most effective prevention measure. The HBV vaccine was initially recommended for individuals at high risk of exposure to HBV infection (health care workers, parenteral drug users, household contacts and infants of infected mothers). However, this strategy had little effect on the incidence of new HBV infections. In 1992, the World Health Organization recommended that all countries include hepatitis vaccine in their routine infant immunization program, especially in areas where hepatitis B is endemic (Kane et al., 1998). Two types of hepatitis B vaccine have been widely used; a plasma derived vaccine (consisting of HBsAg particles and small amounts of LHBs and MHBs) and a yeast-derived recombinant vaccine, consisting of the major S gene product. The latter is used more than the former (McAleer et al., 1984); however, both vaccines are equally efficacious (Andre et al., 1989). Additionally, new recombinant vaccines containing both preS and S antigens are developed to circumvent the non-response to conventional vaccines (Clements et al., 1994; Jones et al., 1998).

Passive immunoprophylaxis is mainly used in the following situations with vaccination: in neonates born to HBeAg-positive mothers (Reesink et al., 1979); after needle-stick exposure (Grady et al., 1978); after sexual exposure (Perrillo et al., 1984); and after liver transplantation (Müller et al., 1991; Samuel et al., 1991).

1.4.2 Treatment of HBV infection

Although mass vaccination programs have begun to control the spread of HBV, therapeutic intervention is the only option for those with established chronic HBV-associated liver disease. Acute HBV does not require specific treatment because more than 90% of adults will spontaneously clear their infection, although symptomatic treatment may be indicated (Gitlin, 1997). The primary objective of therapy for

chronic HBV is eradication of the virus that will, in turn, lead to remission of necro-inflammatory liver disease and an improved long-term prognosis. Two therapeutic approaches have been used to prevent HBV replication: immune modulators, such as interferon (IFN) alpha, and antiviral agents in the form of nucleoside analogues (e.g., lamivudine).

1.4.2A Interferon

Interferons have immunomodulatory, but also antiproliferative and antiviral effects (Goodbourn et al., 2000; reviewed by Karayiannis, 2003). The aim of therapy with IFN alpha is to stimulate the immune system to attack HBV-infected hepatocytes, thereby inhibiting viral protein synthesis. In about 30-40% of chronic HBV patients, IFN alpha therapy will induce a long term remission which is identified by the loss of HBeAg and HBV DNA, normalization of serum aminotransferase levels and improvement in liver lesions (Perrillo et al., 1993; Gitlin, 1997). Indeed, IFN alpha therapy improves the clinical outcome even in the presence of cirrhosis (Niederau et al., 1996). It is also recommended to use IFN alpha therapy in children who have chronic hepatitis after being well selected, on a similar basis to that used in adults, as it may reduce the chance for the development of major complications (Roberts, 2000). Factors that have been associated with a favorable outcome following IFN alpha treatment include high pre-treatment ALT and low serum HBV-DNA levels (Brook et al., 1989; Perrillo et al., 1990; Gitlin, 1997; Nair and Perrillo, 2001). In contrast, factors that have been associated with poor response include male sex, length of chronic state, Asian origin, precore mutations, homosexuality and HIV co-infection (Lai et al., 1987; Lok et al., 1988, 1995; Gitlin, 1997; reviewed by Karayiannis, 2003). Moreover, IFN alpha treatment is expensive, administered by injection and

poorly tolerated with side effects including flu-like symptoms, injection-site reactions, anorexia, rash, neutropenia and thyroid disorders (Vail et al., 1997, Gitlin, 1997).

1.4.2B Nucleoside analogues

The development of nucleoside analogues greatly improved the outcome of HBV treatment. These agents block viral replication directly by being incorporated into newly synthesized HBV DNA, causing chain termination, but they also inhibit the HBV polymerase, and thus viral replication. To date, lamivudine and adefovir are the only nucleoside analogues to have been approved for the treatment of chronic hepatitis B. A number of other nucleoside analogues have been tested or being evaluated presently against HBV including: famciclovir, entecavir, emtricitabine, ganciclovir and FIAU. Early studies have shown that lamivudine reduces rapidly HBV replication and suppresses HBV DNA to undetectable levels after few weeks of treatment (Dienstag et al., 1995; Lai et al., 1997). Furthermore, long term trials in both Asian and Western patients have shown that lamivudine significantly reduced the progression of hepatic histopathological changes and normalized serum ALT levels (Lai et al., 1998; Dienstag et al., 1999; Suzuki et al., 1999; Liaw et al., 2000). Lamivudine has also many advantages over IFN alpha such as better tolerability, oral administration and most importantly, its effectiveness irrespective of the patient's sex or onset of infection (Lai et al., 1998; Maddrey, 2000). However, as with all antiviral therapies, lamivudine is associated with the emergence of strains of HBV that are less sensitive to treatment. The YMDD (tyrosine-methionine-aspartate-aspartate) motif forms the catalytic side of the viral DNA polymerase (see section 1.3.4), and mutations in this and other regions of the gene result in reduced sensitivity to lamivudine (Allen et al., 1998; Lai et al., 1998). The main amino acid substitutions that confer drug resistance are either valine or isoleucine instead of methionine at

position 552 within the YMDD motif (Tipples et al., 1996; Liaw et al., 1999; Chayama et al., 1998).

It has been suggested that combination therapy with two or three nucleoside analogues or combination with IFN alpha treatment might delay or prevent the emergence of viral resistance (Lee, 1997; Rosenberg et al., 1999). Data from one study have demonstrated that the use of combination of lamivudine and IFN alpha is more effective than either monotherapy (Schalm et al., 2000), however, more studies are still needed to confirm the superiority of this combination and to investigate different regimens of combination.

Adefovir, other than acting as a DNA chain terminator, is also thought to stimulate natural killer cell activity and to induce endogenous interferon production (reviewed by Karayiannis, 2003). It is a potent inhibitor of HBV replication (Heijtink et al., 1993; Nicoll et al., 1998). This drug is also active against lamivudine-resistant mutants (Benhamou et al., 2001; Mutimer et al., 2001; Walsh et al., 2001). Adefovir appears to have potential as an effective drug against HBV, and may prove a strong condidate in combination therapies (reviewed by Karayiannis, 2003).

1.4.2C Other Approaches

Immunomodulatory agents such as thymosin α1 (Tα1) whether alone (Mutchnick et al., 1999) or in combination with IFN alpha (Rasi et al., 1996) have been tested in chronic HBV carriers with promising results. However, further trials are still needed to assess this combination.

Therapeutic vaccines are another interesting approach to stimulate the immune system against HBV. As the current HBV vaccines promote production of antibodies and a Th2 biased immune response, for effective therapeutic vaccination, both humoral and CTL responses may be necessary to eradicate infected cells. MF59 and CpG DNA, as

alternative adjuvants showed promising results in human and animals, respectively (Heineman et al., 1999; Wright et al., 1999; Malanchere-Bres et al., 2001; Davis, 2000). Other approaches employing a DNA-based vaccine, which induces both humoral and cellular immune responses, showed promising results in mice, woodchucks and ducks (Geissler et al., 1997; Lu et al., 1999; Rollier et al., 1999). These approaches deserve further studies and clinical trials.

Finally, molecular approaches such as antisense oligodeoxynucleotides (ODN) and ribozymes have been tested in vitro (Nakazono et al., 1996; Moriya et al., 1996; Von Weizsacker et al., 1995; Beck and Nassal, 1995; Welch et al., 1997; Kim et al., 1999); they were efficient in inhibiting viral protein expression and viral replication.

1.4.2D Liver transplantation

Liver transplantation is often the only therapeutic option for patients with acute (fulminant) or chronic liver failure caused by HBV infection (Perrillo and Mason, 1993). However, HBV reinfection of the liver graft is a major problem in these patients who receive immunosuppressive medication to prevent graft rejection. In theses cases, retransplantation may be required due to the rapidly progressive course of the disease which often leads to graft failure and high mortality rates (O'Grady et al., 1992; Samuel et al., 1993). Therefore, prevention of HBV reinfection after liver transplantation is quite important and mainly entails the use of anti-HBs, although escape mutants with frequent mutations within the S gene, especially in the MHR, are often observed in some of these patients (Carman et al., 1996; Protzer-Knolle et al., 1998).

1.5 HBV GENOTYPES

HBV genomes can be classified into at least eight genome groups (genotypes A to H) based on the degree of similarity in their nucleotide sequence. A genetic classification of HBV strains using the nucleotide sequence of the complete genome was performed on 18 HBV clones by Okamoto et al (1988). This classification, based on nucleotide divergence of 8% or more between the strains, enabled the identification of genotypes A to F (Okamoto et al., 1988; Norder et al., 1992a). Genotype G showed a nucleotide divergence of >11.8% from other HBV isolates (Stuyver et al., 2000). Finally, genotype H, which was recognised previously as a divergent clade within genotype F, showed 13.2-15.7% divergence from other HBV strains (Arauz-Ruiz et al., 2002).

Variation within a sub-component of the S gene within the major hydrophilic region (MHR) of HBsAg, the "a determinant", is strongly associated with subtype variation (Okamoto et al., 1987). The "a" determinant is common but there are two pairs of mutually exclusive determinants, d/y and w/r, found at residues 122 and 160, respectively. These characterise four HBsAg subtypes. Subdeterminants of w (w1 to w4) have allowed the further characterisation of ayw and adw subtypes into a total of 10 subtypes: ayw1, ayw2, ayw3, ayw4, ayr, adw2, adw4, adw4q-, adrq+ and adrq-. The q determinants were found to be absent from adr strains in the Pacific region, thus, defining adr and adw as either q+ or q- (Couroucé-Pauty et al., 1978). Collectively, these results indicated that, classifying HBV genomes on the basis of S gene alone is feasible and is consistent with the classification based on the complete genome. Genotypes and subtypes probably evolved in ethnic backgrounds over centuries (Esteban et al., 1999), and the relationship of these ten subtypes to genomic

Table 1.2 Relationship between HBV genotypes and subtypes.

Genotype	A	B	C	D	E	F	G	H
Associated Subtype	adw2 ayw1	adw2 ayw1	adr adrq- ayr	ayw2 ayw3 ayw4	ayw4 adw2	adw4q-	adw2	adw3

groups A to H has been established. Some subtypes can be found in more than one genotype (Table 1.2); hence, confer additional heterogeneity within the genotypes . Strains specifying adw are found in types A, B, E, G, F and H and those specifying ayw in types A, B, D, and E (Norder et al., 1992a; Arauz-Ruiz et al., 2002). Strains specifying r have so far only been found in group C (Okamoto et al., 1988). Between strains, ayw and adw subtypes demonstrate considerable genetic heterogeneity (Norder et al., 1992a). Over the last decades however, subtype determination has gradually been replaced by genotyping.

1.5.1 HBV genotypes: distribution

HBV genotypes have a characteristic geographic distribution, largely in agreement with subtype distribution (Norder et al., 1993a). The HBV subtypes located to smaller and more distinct geographical areas than genotypes (see below). The genotypes of HBV correlated well with geography (Figure 1.11) (Norder et al., 1993b). Genotype A, which might represent a more ancient genomic group (Norder et al., 1993b), is widely distributed in Western Europe, USA, Sub-Saharan Africa as well as Asia and as far as east as the Philippines. Genotypes B and C belong to the indigenous population of South East Asia (Okamoto et al., 1988; Kidd-Ljunggren et al., 1995). In addition, genotype C is found in the populations of the Pacific Islands, and can be sub-differentiated geographically by subtype in this region. As compared to strains

from South East Asia, four adrq- strains formed a distinct cluster within the Pacific area (Norder et al., 1993b). Genotype D has been found worldwide with its highest prevalence in the Mediterranean area, the Middle East, and South Asia (Norder et al., 1993b). Moreover, genotype D has been reported as the only prevalent genotype infecting IV drug users in the Western World (Kidd-Ljungrren et al., 2002). Genotype E is found in West and South Africa. Genotype F demonstrated the highest degree of divergence from the other HBV genotypes (Norder et al., 1993) and is believed to be the original genotype of the New World (Kidd-Ljunggren et al., 2002); it is found in South and Central America. Genotype G has been found in France and the USA (Stuyver et al., 2000) but has not been widely studied. The distribution of genotype H seems to be restricted so far to the Northern part of Latin America including Central America and Mexico (Arauz- Ruiz et al., 2002). HBV strains within each genomic group show a characteristic geographic distribution, and it has been proposed that HBV diverged into genomic groups according to the distribution of humans among the different continents (Telenta et al., 1997).

It is believed that the prevalence of different HBV genotypes in many countries reflects the origin of the immigrants and other patterns of migration (Kidd-Ljunggren et al., 2002) (see sections 1.8.3 and 1.8.4). This is best exemplified by HBV-infected populations living in South Pacific Islands, where the prevalence rate of HBV is amongst the highest in the world. The common genotype is C. This correlates with immigration from South East Asia where genotype C is common (Diamond et al., 1988; Hagelberg et al., 1993; Redd et al., 1995; Kayser et al., 2000) The presence of genotype D in Central America reflects the westward immigration of the Spanish in whom genotypes A and D are common (Arauz-Ruiz et al., 1997).

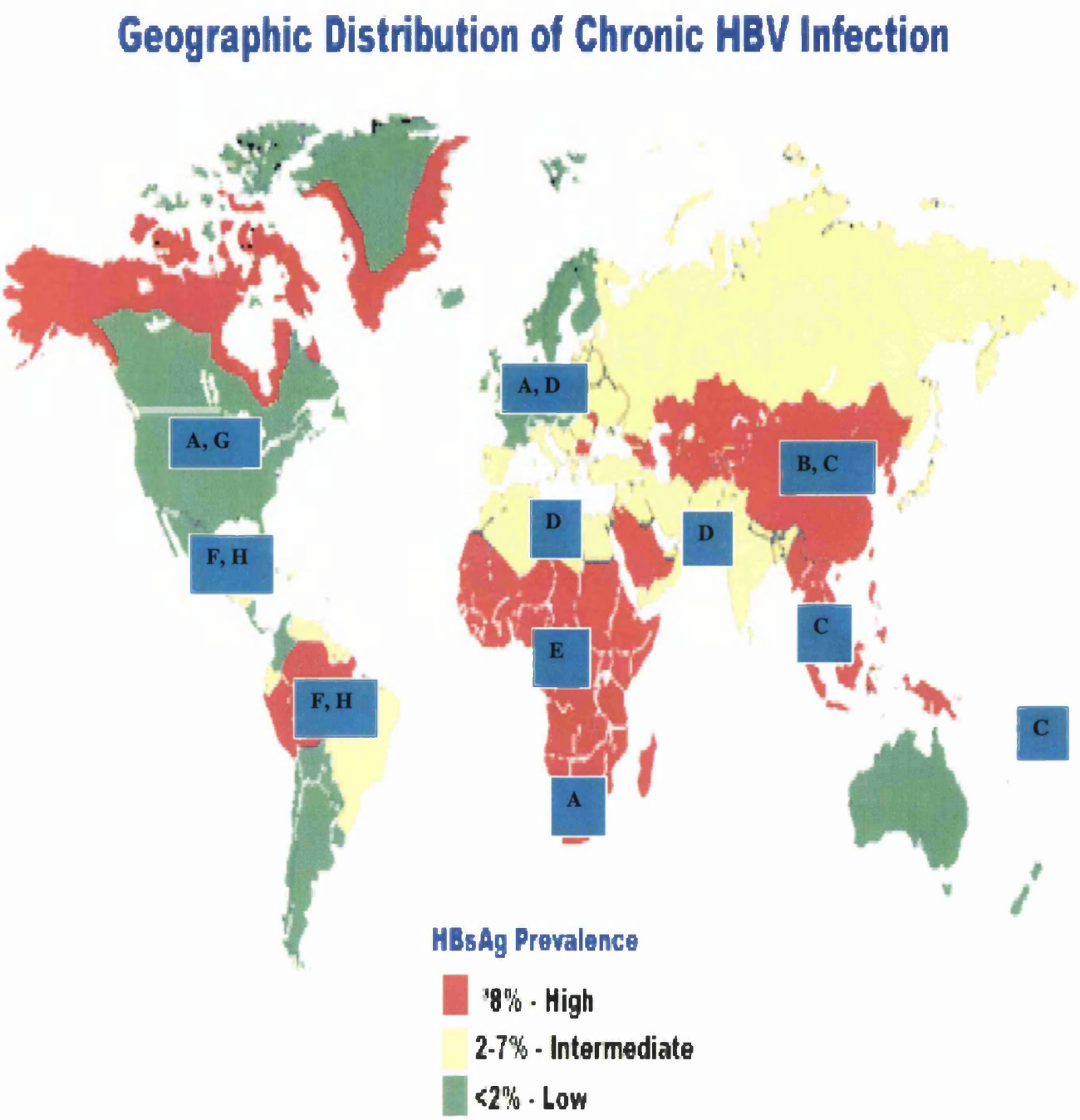


Figure 1.11: HBV prevalence and HBV genotypes geographic distribution.

Further, Yamashita et al. analysed the geographical interface of adr (genotype C) and adw (genotypes A and/or B) strains from Japan. They found an apparent south-to-north gradient of the r determinant and concluded that Japan may have been inhabited originally by people infected with adw (Philippines and Indonesia), and later, people from the Asian continent infected with adr (South East Asia) (Yamashita et al., 1975).

1.5.2 HBV genotypes: clinical relevance

Although the serologic and genotypic classification of HBV has been well documented, the significance of HBV genotypes in terms of the therapeutic response to antiviral therapy remains largely unknown, as does the impact of genotypes on clinical outcome. Some believe that the genotypes do not have any replication advantage over each other or different pathogenic potential. However, many authors have tried to explain the outcome of HBV infection by differences in HBV strains (Shiina et al., 1991; Rodriguez-Fraise et al., 1995; Kao et al., 2000a, b; Orito et al., 2001). So far, these variations in different HBV genotypes have been studied on Far East Asian patients with respect to genotypes B and C as well as New World and Mediterranean patients with respect to genotypes A and D. One example is the prevalence of A1896 in genotypes B, C, and D and its absence in other genotypes, which might explain why this mutation is uncommon among carriers of Western Europe and North America (see section 1.7.3B).

A correlation has been observed between pre-S2 variation and HLA type in chronic patients with high levels of ALT. In a majority of adr- infected patients, the HLA class-I phenotype was A24; in adw and/or ayw it was A2. However, in patients with normal ALT, none of those phenotypes of HLA class-I have been found. Cupps et al. investigated the immune response to an HBV vaccine containing pre-S and S (adw subtype). The ayw peptides from pre-S2 (146-174), which were not present in the pre-

S2+S vaccine, failed to stimulate a T cell proliferation response (Cupp et al., 1993). This is supported by the observation that the T- and B cell response is mostly subtype-specific. In mice immunized with peptides containing the pre-S2 region of the d and y subtypes, Milich and co-workers found that the responses were pre-S2-subtype-specific. They suggested that any new generation of HB vaccines should contain both the d and y subtype sequences of the pre-S region to increase the S-specific antibody response (Milich et al., 1990).

In several studies, a link between HBV genotype and risk of chronic outcome has been reported: an association of genotypes B and C with chronicity was reported in South East Asian studies (Ogawa et al., 2002). A large proportion of such HBeAg-positive patients were infected with subtype adr, which might to some extent be related to the fact that the pre-C stop codon mutation is more common in genotype B (Shiina et al., 1991; Lindh et al., 1999; Yuen et al., 2003). Genotype B may be associated with the development of HCC in young HBsAg carriers, suggesting possible pathogenic differences among HBV genotypes (Kao et al., 2000 a,b; Nakayoshi et al., 2003). On the other hand, core promoter mutants have been reported more commonly in patients with genotype C than with genotype B (Orito et al., 2001; Yoo et al., 2003; Yuen et al., 2003; Kao et al., 2003). Genotype C-infected patients often have more advanced liver disease with more acute exacerbations (including HCC) compared to genotype B (Kao et al., 2000a, 2001, 2003; Chan et al., 2003; Nakayoshi et al., 2003). Similarly, genotype A was suggested to lead more often to chronicity as it was found more often in chronic HB patients than genotype D, whereas, the opposite situation was found in patients with acute hepatitis (Mayerat et al., 1999). Sustained remission after seroconversion to anti-HBe was reported to be higher in genotype A than genotype D or F (Sanchez-Tapias et al., 2002). Genotype D

is associated with more severe disease in post-transplant patients (in whom the pre- and post-transplant sequences were similar) (McMillan et al., 1996). Further, genotype D may be associated with fulminant hepatitis more commonly than predicted from its prevalence in chronic carriers. Similar to genotype B, most genotype D strains in anti-HBe positive patients have pre-C mutations (A1896 and/or A1899) (Sanchez-Tapias et al., 2002).

Recombination between genotypes has been reported mainly from South-East Asia. Recombination between genotype A and C from Vietnam (Hannoun et al., 2000b), between C and D from Tibet (Cui et al., 2002), and between B and C from South-East Asia (with the exception of Japan) (Sugauchi et al., 2002, 2003) have been reported. Using nucleotide sequences from databases, Morozov and colleagues analysed the evolution of complete HBV genome and three ORF sequences: S, C and X. They found hot spots for recombination in those ORFs. As there is a high recombination rate between genotype B/C, they suggested that such genomes could obscure the real genotype-related differences in the pathogenicity of HBV (Morozov et al., 2000).

Genotypic changes have been reported in several studies after anti-HBe seroconversion: genotypic shift from D to A (Friedt et al., 1999) and A to D (Gerner et al., 1998). However, it seems unlikely that a new genotype could evolve under selection within such a short period a time. These observations may indicate that infection with different genotypes of HBV can occur simultaneously (although there is no direct evidence) (Gerner et al., 1998; Morozov et al., 2000).

There are hints that genotype C may be more resistant to IFN (Kao et al., 2000b; Wai et al., 2002). This is similar to hepatitis C, where genotypes 2 and 3 are more sensitive to IFN therapy than 1 or 4 (Hoofnagle et al., 1997). This finding might be related to the fact that both core promoter and pre-core stop codon mutations have been reported

to influence response to IFN (Erhardt et al., 2000). In general, it has been described that Asian HBV carriers have a lower response rate to IFN than non-Asian carriers (because of infection at an early age) (Lai et al., 1987; Lok et al., 1988; Hoofnagle et al., 1990), in keeping with the different genotypes infecting Asian and non-Asian ethnic groups. It may also be that the adw subtype is associated with a higher risk of lamividine-resistance than ayw (Zollner et al., 2001, 2002). However, this observation was in contrast to the finding that genotype B has a better virologic response to lamivudine compared to genotype C and that both genotypes showed a similar risk of developing lamivudine-resistance after 1 year of therapy (Kao et al., 2002). Taken together, these data suggest pathogenic and therapeutic differences among HBV genotypes and indicate a potential role for HBV genotypes in the virus-host relationship.

1.5.3 HBV genotypes: HBcAg amino acid variability

The genetic differences between HBV genotypes are not limited to the S gene. Characterisation of HBV strains by subtyping and/or genotyping using other HBV proteins has been used for tracing a source of infection (Ogata et al., 1993; Hawkins et al., 1996; Kidd-Ljunggren et al., 1999)

Although the amino acid sequence of the core region is relatively conserved (compared to surface gene), amino acid substitutions have been observed in several reports (Ono et al., 1983; Okamoto et al., 1986; Ferrari et al., 1991; Ehata et al., 1992,1993; Chuang et al., 1993; Akarca et al., 1995a; Hosono et al., 1995; Hur et al., 1996; Alexopoulou et al., 1997; Carman et al., 1997b; Fujiwara et al., 1998; Gunther et al., 1998; Takahashi et al., 1998) (Table 1.3). This is largely linked to the clinical or serological picture (Chuang et al 1993; Ehata et al, 1992,1993; Akarca et al., 1995a; Carman et al., 1995b). Core variability has been shown to correlate with disease

activity. Mutations within the core gene have been linked to persistence (Carman et al., 1995b; Ehata et al., 1992,1993; Akarca et al., 1995a; Hosono et al., 1995). Although core aa substitution was reported in many studies, some of these variants are actually the consensus sequence from other genotypes. Therefore, in view of the heterogeneity of HBV strains, it is difficult to determine whether substitutions recorded only once represent naturally occurring variants or merely an immune-selected mutation in those patients. For example, based on the results presented in chapter 3.1.2, aa residue 74 of the core gene contains 5 possible natural aa variations in both HBeAg and anti-HBe positive carriers, indicating that these are normal variations linked to different ethnic/HBV genomic groups; in some studies however, these variations were reported as true mutations after HBeAg seroconversion, IFN therapy etc. Perhaps this reflects the variable definition of a "standard" sequence in such studies, due to different genotypes, as well as differences in ethnic background. So far, no systematic study has been conducted to correlate HBcAg diversity with HBV genotypes and/or subtypes and ethnic background. This background information would not only be critical for the correct interpretation of sequences observed in clinical studies, but may aid in the design of immunotherapies specific for persons of a particular ethnic origin. Until now, no identical C gene consensus sequence has been identified due to its heterogeneity between different HB genomic groups and subgroups in different geographic regions. The published core aa sequences in the databases therefore do not represent the consensus aa/nucleotides occurring at critical positions, and conserved HB core protein sequences are rarely reported. There are a number of issues to be considered in developing a C gene database. First, the core region is highly divergent in anti-HBe carriers (Sanantonio et al., 1991 a, b; Carman et al., 1995b). Consequently, core sequences obtained during

the HBeAg-positive phase should have less disease-related sequence deviation. Second, as the most conserved regions of the C gene are short stretches in approximately the first and the last 150 nucleotides (Kodama et al., 1985; reviewed by Lok, 1997), the middle section should be used for this purpose. Third, studies have shown that in certain populations where HBV is endemic, a higher variability of HBV might be expected (Carman et al., 1997b). Thus, T-cell epitopes could have various locations on account of the divergent distributions of HLA types in different countries and different aa sequences in the HBV subtypes (Chuang et al., 1993) (see section 1.7.1). Consequently, it is important to sequence geographically divergent HBV genomes to define the extent of natural variation (Norder et al., 1994).

Table 1.3: The most frequent amino acid changes in HBcAg sequence (hot spot positions).

Amino Acid Position	Amino Acid Change
5	T, H, Q, A, S,P
12	A, P, Y, S, T
13	A, L, T, E, M, G, V
21	A, D, H, F, N, V, P, S, T
27	A, L, G,N, S, T, I, V
35	A, S, T
38	F, H, G, N, Y
40	A, D, E, K, Q
49	C, E, S, T, Y
50	A, H, N, P, S, T
55	I, L
59	G, I, V, S, T
60	E, I, L, V
63	A, E, D, G, N, P, R, V
64	D, E
67	G, S, N, T
74	A, F, S, I, R, G, V, T, N
77	D, , E, Q, T
79	P, Q
80	A, G, I, S, T, V
84	A, D, L, Q, R, V
87	G, T, S, N
91	I, L

92	H, T, N
93	L, M, T, V
97	I, F, L
100	I, F, L, T
105	I, L, T, V
113	D, E, L, Q, S, V
116	I, L, V
130	A, L, N, P, Q, S, T
135	A, E, L, P, Q, S, T
147	A, C, S, T
151	C, E, G, P, R, Q
153	C, D, G, R, V
155	P, S, T, W
181	P, S

Note: amino acids numbered from the beginning of HBcAg. The normal variations are on the basis of different genotypes studied.

1.6 HBcAg EPITOPES

The identification of immunodominant T cell epitopes within the core molecule could theoretically be useful for the design of more effective alternative vaccines against HBV infection and possibly to plan future strategies to manipulate the immune response to HBV in subjects who do not spontaneously clear the virus. The analysis of antigenic determinants has been successfully accomplished by monoclonal antibodies. Since high titres of anti-HBc antibodies are produced to a range of epitopes during the course of the infection, this has allowed more accurate location of the epitopes. However, epitope mapping of viral protein sequences with mAbs does not always identify epitopes identical to those recognized by human B cells, despite their ability to specifically compete with human antibodies. Also, different mAbs with identical specificity in cross-competitive experiments may give varying results when mapping epitopes (Sallberg et al., 1991). A summary of different HBcAg immune epitopes is shown in figure 1.12 and table 1.4.

1.6.1 B cell epitopes

1.6.1 A HBcAg B cell Epitopes

The region of HBcAg involved in binding to naïve B cells has not been fully mapped (Lazdina et al., 2003). However, studies in humans and mice have shown that the binding of HBcAg to naïve B cells occurs between linear motifs on the surface of immunoglobulin molecules (Lazdina et al., 2001) and the repetitive, highly ordered protein spikes on the core shell which might be optimal for cross-linking of B cell antigen receptors (Bottcher et al., 1997; Milich et al., 1997a). HBcAg is thought to form a highly conformational epitope, which is not recognized when denatured by detergents. On the other hand, the HBeAg epitopes are presumed to be linear, since HBcAg denaturation does not prevent the binding of anti-HBe to the detergent-treated

HBc particles (Ferns and Tedder, 1984). Further, the capacity of anti-HBc mAb to inhibit the binding of human polyclonal anti-HBc indicated that the majority of anti-HBc bind to one area of the molecule and suggested the existence of either a unique or immunodominant epitope (Ferns and Tedder, 1986; Waters et al., 1986). In cross-competition assays mAbs recognized a broadly similar antigenic determinant on HBcAg (even in the presence of human antisera). It was therefore proposed that HBcAg, unlike HBeAg, has a single reactivity (Ferns and Tedder, 1986; Waters et al., 1986). However, Takahashi et al., (1983) obtained mAbs against two determinants that are expressed on the surface of the core particle. Using pre-S as a foreign epitope inserted at different regions of HBcAg, Schodel et al, (1992) found that insertion between residues 75 to 83 abrogated recognition of HBcAg by 5 of 6 anti-HBc monoclonals and diminished recognition by human polyclonal anti-HBc.

The exact location of the HBcAg dominant B cell epitope remains controversial after studies using synthetic peptides and monoclonal as well as polyclonal antibodies. However, it appears to be between aa residues 73 to 89 (GVNLEDPASRDLVVSYV) (Salfeld et al., 1989; Tordjeman et al., 1993; Pushko et al., 1994; Rodriguez-Frias et al., 1995). In an attempt to localise immunodominant epitopes with a different approach, Conway and colleagues (1998) used cryoelectronmicroscopy to pinpoint the binding site of a monoclonal antibody that recognizes residues 78 to 83. Residues 73-83 seem to form a non-helical loop between two alpha-helical segments at the tip of the capsid spike (Bottcher et al., 1997; Conway et al., 1998b). The second B cell epitope, residues 128-135 of the HBc protein, contains a shared T cell (Milich et al., 1989), and B cell determinant, and P129 has been found to be essential for mAb binding to the B cell epitope. Thus, this is an example of a T cell-dependent B cell epitope (Sallberg et al., 1991). HBeAg, however, has been proposed to be a strictly T

cell dependent Ag (Milich et al., 1988). Another immune epitope has been located between aa residues 107-118, though this is linear (Colucci et al., 1988; Sallberg et al., 1994). Finally, another group of antigenic and immunogenic sites have been identified within the C-terminus region of HBcAg (aa 148-160), called HBicAg (hepatitis B inner core Ag) (Machida et al., 1989; Sallberg et al., 1994). Only one of these epitopes (amino acid residues 107-118) has a relatively conserved sequence; the other two have a high rate of variation. Collectively, in these reports, anti-HBc recognized many epitopes located in the HBcAg. This indicated that the antigenicity of the core particle is due to many epitopes that are distributed over the entire primary structure of the HBc protein, although once they are assembled into the capsid, they are unable to bind more that one antibody molecule, because of steric hindrance (Tordjeman et al., 1993).

1.6.1B HBeAg B cell epitopes

Because HBcAg and HBeAg share an identical 149 aa sequence, there is a strong possibility that the two antigens may display cross-reactivity. Peptide residues 74-89 are recognized by both anti-HBc (Tordjeman et al., 1993) and anti-HBe-specific antibodies (Sallberg et al., 1993).

There are two major HBeAg B cell epitopes. HBc/HBe1, is a linear region spanning aa 76-89 (Ferns and Tedder, 1984; Salfeld et al., 1989). This is on the surface of the core particle and overlaps with the conformational HBc determinant. At the C-terminus of HBeAg, aa 130-138 form the HBe-2 epitope (Ferns and Tedder, 1984; Salfeld et al., 1989; Sallberg et al., 1993). Unlike the HBc epitope, HBe-1 was found to be linear (Waters et al., 1986; Salfeld et al., 1989). This would imply the concomitant presentation of HBc/HBe antigenic determinants in two major structural states: in a dimeric, HBeAg state, exposing two antigenic determinants (the HBe-2

determinants and the HBc/HBe-1 determinants as HBe-1 only); or, in the core particle, HBe-2 is masked and additional conformational HBc epitopes are introduced into the exposed HBc/HBe-1 determinant (Salfeld et al., 1989). This would imply a shared antigenic region between two Ags encompassing different regions in several reports. That HBcAg and HBeAg share structural features is supported further by Salfeld's finding that the aa sequence around position 80 is exposed and recognised as a major antigenic determinant in both HBcAg and HBeAg; a surface location of this sequence is also suggested by its preferential tolerance to naturally occurring mutational aa changes. Unlike the HBe-1 epitope, HBe-2 has little aa variation within HBV subtypes. One of the reasons might be that the two prolines at residues 129-130 are essential for recognition by most mAbs, perhaps for steric reasons (the C-terminal part of HBeAg may be internal and thereby less accessible to antibodies) (Sallberg et al., 1991, 1993).

1.6.2 T Helper Epitopes

Ferrari and colleagues described a dominant T-helper cell epitope (positions 50-69 in HBcAg) which is recognised by CD4+ lymphocytes from 95% of patients with active or chronic HBV infection. They identified two additional less frequently recognised sequences in the majority of the patients: residues 1-20 (in 69%) and 117-131 (in 73%) of patients regardless of their HLA backgrounds (Ferarri et al., 1991; Penna et al., 1992) (Figure 1.12; Table 1.4). Interestingly, when epitope 50-69 is not dominant, T cells are mainly focused on epitopes 1-20 and/or 117-131. From anti-HLA class-II blocking experiments these epitopes preferentially activated $CD4^+$ T cells. These T cell responses occurred irrespective of the patients genetic background, in contrast to Milich's observation that the fine specificity of the murine T cell response to HBcAg was more dependent on the MHC haplotype of the responder strain (Milich et al.,

1987). These dominant T cell epitopes are located in the portion of the nucleoprotein molecule common to both HBcAg and HBeAg, which confirms, as described earlier in this chapter, that core and e Ags are cross-reactive.

1.6.3 CTL Epitopes

Using different synthetic peptides, at least two different HLA class-I-restricted CD8-positive CTL epitopes have been identified so far within the HBcAg, which are both shared by HBc and HBe Ags. First, aa residues 11-27, an HLA-A2-restricted epitope, has been shown to induce a specific CTL response during acute HBV infection (Bertoletti et al., 1991).

Table 1.4: Proposed antigenic epitopes within HbcAg.

Epitope/Amino Acid Positions	Type of Cell/Epitope	Author
1-20	Th	Penna-1992
18-27	CTL	Bertoletti-1991
50-69	Th	Ferrari- 1991
74-89	B	Salfeld- 1989
84-101	CTL	Ehata-1992
107-118	B	Colucci-1988
117-131	Th	Ferrari-1991
128-135	B	Salfeld-1989
141-151	CTL	Missale-1993
148-160	B	Machida-1989

Note: Numbers indicate amino acid residues within the core protein. B, Th and CTL represent B-cell, T helper and CTL epitopes, respectively.

Second, aa residues 141-151 have been shown to induce a specific CTL response that was restricted by two independent HLA class-I alleles in vivo in human; HLA-31 and HLA-Aw68 (Missale et al., 1993). The latter responses have been shown to focus on precisely the same 11-residue sequence, recognized by patients with acute HBV.

From various clinically-based observations, a number of CTL epitopes have been proposed within the core gene encompassing aa residues 84-101 (Figure 1.12, table 1.4) (Ehata et al., 1991, 1992, 1994); 141-151 (Khakoo et al., 2000) and 147-155 (Chuang et al., 1993). The basis for identification of such CTL epitopes was that chronically HB-infected patients with ineffective CTL responses showed a relatively high mutation rate in these regions. They therefore suggested that these putative CTL epitopes might be sites under immune pressure, leading to emergence of CTL- escape mutants. However, they did not present data to support the existence of such CTL epitopes.

1.7 HBV PATHOGENESIS

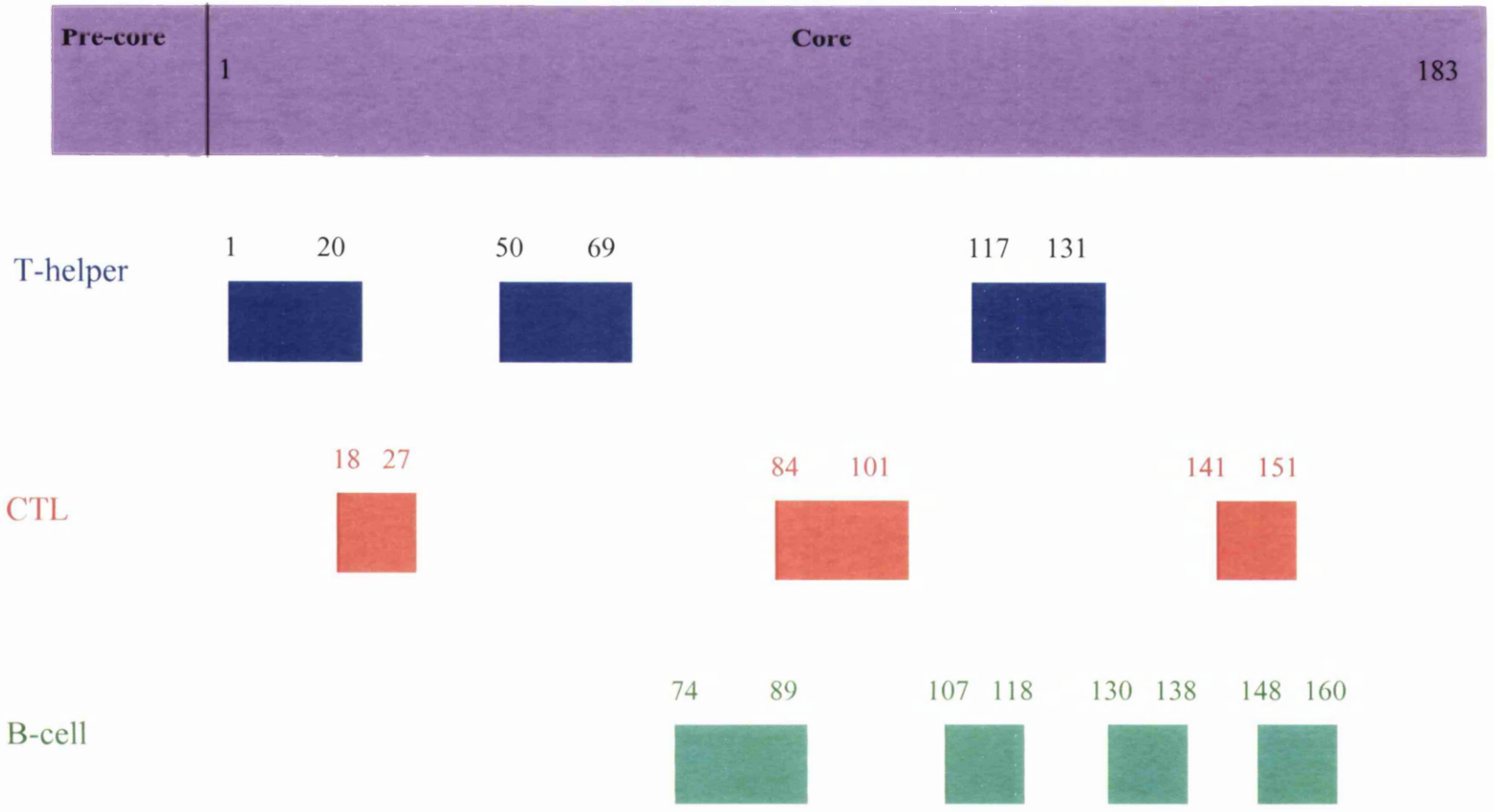
The diversity of clinical syndromes and disease manifestations associated with HBV strongly suggest that the outcome of this infection is determined by the quality and vigour of the antiviral immune response produced by each infected individual. Several factors must coexist for the immune response to clear the virus. First, the local immune response must be vigorous enough to mount a strong CTL response. Second, the infected cell must be able to produce the appropriate antiviral activity in response to the immune cell-mediator. Third, the virus life cycle must include functions that are susceptible to those antiviral factors, e.g., degradation of viral RNA post-transcriptionally by CTL and inflammatory cytokines (Tsui et al., 1995; Guidotti et al., 1997). The pathogenic mechanisms responsible for liver cell injury in HBV infection are not well understood, though it appears that the virus is not directly

cytopathic for the infected hepatocytes. This finding is based on one of the observations that ASC (asymptomatic chronic) HBV carriers have normal liver morphology and function despite the presence of high levels of viral replication within their hepatocytes (Milich et al., 1995a). In contrast, a direct cytotoxic effect of large amounts of accumulating viral proteins could cause a sudden necrosis of the hepatocytes (Sterneck et al., 1997). Chisari et al (1987) has shown in a transgenic mouse model that intracellular accumulation of large surface protein is directly cytopathic to the hepatocytes (ground glass appearance). Similarly, introduction of a pre-S mutant (G133E) into ducklings resulted in hepatic destruction (Lenhoff et al., 1998). High level accumulation of intracellular HBcAg can also be cytotoxic for hepatic cells (Roingeard et al., 1990a; Rosmorduc et al., 1994).

1.7.1 HLA Antigens and HBV pathogenesis

HLA associations with HBV infection have provided evidence for a genetic predisposition to development of the chronic carrier state. Antigen-specific T cell receptors recognise peptide bound to HLA molecules and the repertoire of peptides bound depends on the HLA type of the individual. Therefore, HLA type is likely to influence susceptibility to chronic HBV infection. Although each individual HLA allele has the potential to bind many different peptides, there is a degree of specificity that results in different HLA molecules having individual peptide-binding repertoires. This is widely thought to be the molecular basis for many of the well known associations between HLA type and chronic infections, such as malaria (Almarri et al., 1994). The relationship between HLA types and certain diseases has focused attention on HBV persistence and, in particular, DR alleles. The hypothesis that the outcome of HBV infection is influenced by the HLA-associated immune response has

Figure 1.12: Potential immunodominant domains of HBcAg.



Note: Numbers indicate amino acid residues.

been evaluated in many studies. For example, the association of HLA-BW15, HLA-A1 and B8 with persistent HBs antigenemia was known in the 1970s (Hillis et al., 1977). HLA-B7 in white carriers (Sampliner et al., 1981), but DQA1*0501, DQB1*301 and DRB1*1102 in African-American carriers (Thio et al., 1999) were found to confer resistance to chronicity. Also, an association between resistance to chronicity and HLA-DQB1*1101/1104 and HLA-DQA1*301 has been reported in Chinese patients (Jiang et al., 2003). Carbonava et al, (1983) demonstrated a linkage between HBV chronicity and HLA-B40 in males as well as AW30, CW5, B18, B14 and DR1 in female patients. Moreover, they found a negative association with DR4 in females, in keeping with a hypothesis that even if HLA is not of major importance in the carrier state, a protective effect by DR4 may exist, similar to DR2 in type I diabetes. There is an association between a self-limited course of acute HB and presence of HLA-DR B1-1302 in West Africa (Thursz et al., 1995), DR B1-1301 and 1302 among Caucasians (Hohler et al., 1997) and DR13A in Asians (Ahn et al., 2000). This indicates a strong association between HLA-DR13 and more vigorous HBcAg-specific $CD4^+$ T cell responses, which are associated with a self-limited course of HBV infection and protection from chronicity (Diepolder et al., 1998; Cao et al., 2002). DR B1-1301 and 1302 are also believed to protect against non-responsiveness to the recombinant HBV vaccine (Hohler et al., 1997, 2002). In keeping with these findings, alleles DRB1*01, DRB1*I1, DRB1*15, DQB1*0501 and DPBI*0401 were strongly associated with a higher response (reviewed by Milich, 2003). Furthermore, immunodominant CTL epitopes restricted by HLA-A2404 have been identified in East Asia (Sobao et al., 2001). Collectively, these results suggest that the epitopes for CTL/Th recognition might be different on account of the diverse distributions of HLA antigens in different geographic regions. Analysis of patients

with different HLA haplotypes revealed that several sequences within the HBV proteins, in particular core, could induce significant levels of T cell response. These issues will be discussed in more detail below.

The presence of HBcAg-specific-$CD4^+$ and $CD8^+$ T cells in the infected liver is well documented (Ferarri et al., 1987; Penna et al., 1992). The relation between HLA class I and II restricted T cell responses to nucleocapsid antigens in patients at different stages of HBV infection has been studied in several clinically-based studies. Various immunodominant epitopes for both T cell groups have been identified.

1.7.1A HLA-class I Molecules

Previous results, mostly based on CTL studies, have shown that HLA class I antigens have been involved as target structures in viral clearance. Using anti-HLA monoclonal antibodies, Lakhdar et al and Wallace et al, blocked the lysis of an EBV-transformed cell line by EBV-specific CTLs reactivated in vitro (Wallace et al., 1981; Lakhdar et al., 1984). Similar observations have been noted for influenza virus (McMichael et al., 1977).

The presence of HLA class-I and class II on the hepatocyte membrane has been described by Montano et al (1982) and Lautenschlager et al (1984). They found that display of HLA class I in chronic HBV infection was closely related to both the HBeAg/anti-HBe status and histological activity; HLA class I Ag was undetectable on the hepatocyte membrane in the early immune tolerance phase of chronic HBV infection. These Ags became detectable in the immune clearance phase and persisted or even became more intense after seroconversion when replication ceased and inflammatory activity became minimal. They further suggested that the increase in expression of HLA class I in chronic HBV infection probably occurs during the period of seroconversion. Thus, expression of the viral determinants on the membrane

of the hepatocytes seems closely related to active intrahepatic HBV replication (Ray et al., 1976; Chu et al., 1988). Similarly, extending their previous observations on mice, McMichael et al observed that CTL recognise viral Ags on hepatocytes only in the context of HLA class-I (McMichael et al., 1977).

HLA-A2-restricted epitopes have been mostly used for studies of HBV-specific CTL in patients with hepatitis B (Bertoletti et al., 1991; Penna et al., 1991; Chuang et al., 1993; Hwang et al., 2002; Shimada et al., 2003), because it is a common haplotype. The most dominant HLA-A2-restricted CTL epitope within the core is from aa 18-27 (Bertoletti et al., 1991). Missale et al, (1993) demonstrated that two independent class I molecules, HLA-A31 and HLA-AW68 bound to core aa 141-151 in acute HBV infection. In a small number of Southeast Asian patients, HLA-A24 was the most common HLA class I allele found in this region (Soboa et al., 2001).

In several studies in transgenic mice, CTL have also been shown to clear HBV RNAs, proteins, nucleocapsid particles, and DNA replicative intermediates from the liver and virions from the serum of these animals, mostly by non-cytopathic mechanisms (Guidotti et al., 1999, 2001, 2002; Kakimi et al., 2001; Baron et al., 2002; Thimme et al., 2003). The clearance occurs quite rapidly without massive destruction of the liver. Thus, it is possible that the immunopathological and antiviral effects of the CTL response could be due either to the direct cytopathic effect or to the action of antigen-nonspecific mediators, which post-transcriptionally degrade the viral RNA and possibly cccDNA without a cytopathic effect (Tsui et al., 1995; Guidotti et al., 1997, 1999; Kimura et al., 2002; Thimme et al., 2003). This effect of non-cytolytic clearance of HBV from the hepatocytes by CTL-derived cytokines is several orders of magnitude more efficient than its destructive effect (Figure 1.13) (reviewed by Chisari et al., 1997), and it has been shown to decrease total HBV DNA by >50-fold

(Thimme et al., 2003). The potential for a single CTL to kill many infected hepatocytes in vivo may be much less than the cytolytic capacity of the same CTL in vitro, because of differences in both target cell sensitivity and experimental conditions (Guidotti et al., 1996).

Like the case of HBV, other viruses, including adenoviruses (Zhang et al., 1998), coxsackievirus (Horwitz et al., 1999), and measles virus (Parra et al., 1999) are susceptible to the antiviral activity of cytokines produced by CTLs. Moreover, non-cytopathic inhibition of replication of these and other viruses including retroviruses, influenza viruses, vesicular stomatitis virus, HSV, vaccinia virus and reovirus by IFN-α/β has also been reported (Vilček, 1996; Samuel et al., 2001).

1.7.1B HLA class-II Molecules

The humoral immune response during HBV infection is an important indicator of underlying T-cell function or dysfunction (Milich et al., 1995). The main function of HLA class II gene products is to bind antigenic peptides derived from exogenous and endogenous proteins and to present them, at the cell surface, to the T cell receptor on $CD4^+$ T lymphocytes Penna et al., (1992) suggested that HLA class-II restricted $CD4^+$ cells could also potentially participate in the clearance of infected hepatocytes. Expression of HLA class II determinants on the liver cell surface following HBV infection has only been reported by Van Den Oord et al., (1986) which therefore remains to be clarified. Whether $CD4^+$ T cells directly cause liver injury or mediate it through the activation of HBc/eAg-specific $CD8^+$ CTL is controversial. It is possible that $CD4^+$ T cells mediate liver injury directly through a FAS-FASL mechanism, or alternatively, through the production of cytokines by a CD4+ Th_1-like response (see below) (Milich et al., 1997b; Chen et al., 2001; Cao et al., 2001). The T cell independence of HBcAg was based on the observations that naïve human B cells

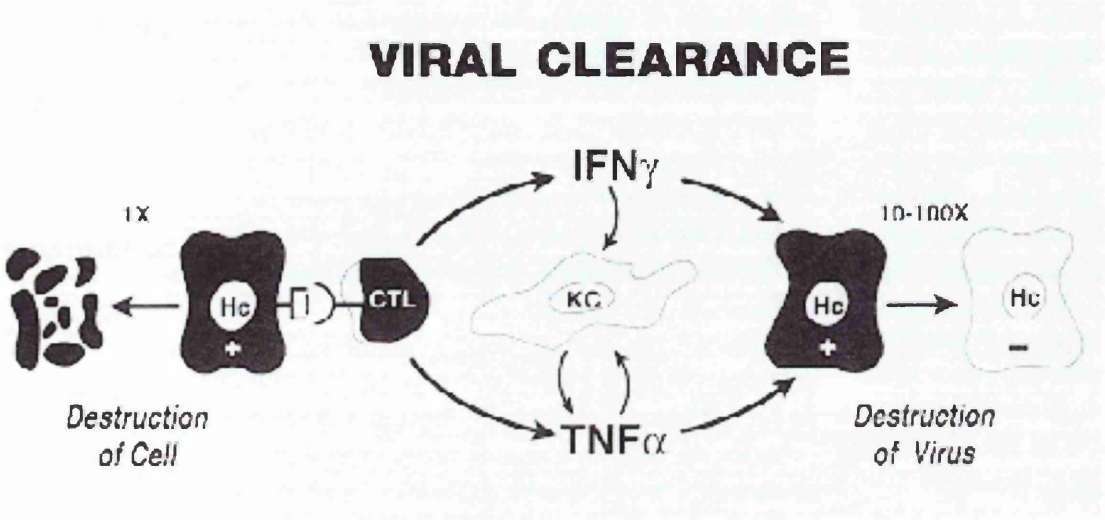


Figure 1.13: Noncytolytic clearance of HBV from the hepatocytes by CTL-derived cytokines. As shown the curative effect of the CTL response is several orders of magnitude more efficient than its destructive effect. Taken from Chisari et al., 1997.

were able to produce HBcAg-binding human IgM; however, HBsAg was unable to induce the production of anti-HBs antibodies in the same manner (Milich et al., 1986; Cao et al., 2001). On the other hand, high levels of antibody production in mice immunised with a mutant HBcAg indicated the ability of the core particle to induce B cell immunogenicity through CD4+ T cell priming as well (Lazdina et al., 2003). The humoral response to HBsAg is T cell dependent and shows great variation with a significant fraction of humans responding poorly. The S gene product is also a relatively poor inducer of the cell-mediated response (Ferarri et al., 1990; Milich et al., 1995a; Reignat et al., 2002).

There are distinct subsets of regulatory CD4+ Th cells designated as Th_1 and Th_2 cells. Mouse experiments have demonstrated that HBcAg-induced cytokine responses

are epitope-dependent. If the linkage of this skewing toward either a Th_1- or Th_2-like response in transgenic mice holds true for man, this would form the basis for linking the T cell subsets to the pathogenesis of chronic disease. In this regard, a number of investigators proposed an important role for cytokines as having both direct anti-HBV properties as well as determining the outcome of HBV infection. This is due to observations that Th-subsets are largely defined by a unique pattern of cytokine secretion.

Allelic differences in HLA molecules can modulate their ability to bind peptides, and thereby change the nature of T cell recognition by Th subsets that have different patterns of cytokine production (Abbas et al., 1996). The effects of immunomodulatory cytokines (IL-12, IFN-α, IFN-γ) on HBV suppression has been studied in some detail (Porres et al., 1988; Fattovich et al., 1989; Kakumu et al., 1991; Guidotti et al., 1994; Brunetto et al., 1995; Aikawa et al., 1995; Rossol et al., 1997; Suri et al., 2001, Hyodo et al., 2003). Following stimulation, HBc/e antigen-specific T cells progress from naïve cells to Th_0 cells that further differentiate into Th_1 or Th_2 cells. HBcAg-induced Th_1 cells mainly produce interleukin IL-2, IL-12, TNF-β and IFN-γ, thereby favouring cellular immune responses such as macrophage activation and delayed-type hypersensitivity (DTH) which can also lyse HBcAg-specific human B cells (Milich et al., 1995b, 1997b; Cao et al., 2001). Th_2 cells mainly produce IL-4, IL-5, IL-6, IL-10 and IL-13, cytokines that not only stimulate humoral responses but can also suppress Th_1 cellular responses (Milich et al., 1997c, 1998; Penna et al., 1997). Th_1 and Th_2-type responses are thus antagonistic. Because both HBcAg and HBeAg are produced during wild-type HBV infection and because the Th_1 and Th_2 subsets are cross-regulatory, a dynamic balance between the HBcAg/HBeAg-specific Th_1 and Th_2 subsets may exist. If so, alterations in the Th_1 -Th_2 cell balance would be

expected to influence the course of the infection in terms of both liver injury, persistence, and antiviral therapy (Millich et al., 1997c; Hultgren et al., 1998, Hyodo et al., 2003).

1.7.2 ACUTE HBV

The consequence of acute HBV infection in adults depends strongly on their immunological status (Koibuchi et al., 2001). During the early stages of acute HBV infection, before the immune system is fully mobilised, the first line of defense appears to be the production of interferon. It may exert an antiviral effect and enhance natural killer cell cytotoxicity.

1.7.2A NKT and NK cells

HBV virions and subviral particles contain glycolipids and phospholipids derived from the cellular membrane (Ganem et al., 1996), that can be processed by $CD1^+$ cells (professional APCs and hepatocytes) and presented to intrahepatic natural killer T (NKT) cells (direct mechanism) (Kakimi et al., 2000). NKT cells are particularly abundant in liver and they are thought to play a role in HBV-immunity by producing anti-viral cytokines. NK cells are promptly activated by NKT cells and enhance induction of IFN-γ production. Since the induction of IFN-γ and IFN-α/β as well as the inhibition of HBV replication occurs before a significant number of inflammatory cells are recruited into the organ, it is likely that these cytokines are produced by cells that reside in the liver.

1.7.2B CTL/Th Response

Enhanced T cell responses directed to HBcAg in patients with acute HBV temporarily coincide with viral clearance. A few class-I and class II-restricted T cell epitopes have been defined in HBV proteins. The envelope and polymerase antigens are more often targeted by the class I-restricted CD8-positive CTL response in such patients

(Nayersina et al., 1993; Chisari and Ferarri 1995; Rehermann et al., 1995). In acute infection an efficient CTL attack against hepatocytes expressing HBcAg-derived peptides plays a major role in the selective lysis of infected cells. This response extends to multiple viral epitopes (Missale et al., 1993); it is polyclonal and multi-specific (reviewed by Chisari and Ferarri, 1995; Nayersina et al., 1993). In an important study, Bertoletti et al (1993) demonstrated that 90% of HLA-A2-positive patients with acute HBV infection produce an HLA-A2-restricted, CD8-positive CTL response to a 10-residue epitope that maps between aa 18 and 27 of HBcAg. Interestingly, this response was not detectable in HLA-A2-positive patients with chronic infection (Penna et al., 1991). Several HLA class I-restricted epitopes, specially core aa 18-27; envelope aa 183-191, aa 250-258 and aa 335-343; and polymerase 455-463 are recognized by the majority of patients during acute HBV (Nayersina et al., 1993; reviewed by Chisari and Ferarri, 1995). Thus, polyclonality and multispecificity of the CTL response mitigate against the emergence of viral escape mutants during acute viral hepatitis (Chisari and Ferarri, reviewed 1995).

HBcAg is a major immune target (Mondelli et al., 1982; Chu et al., 1987; Ferrari et al., 1987; Milich et al., 1989) and serves as an inducer of the Th cells (Marinos et al., 1994). Also, the class II-restricted, CD4-positive helper T cell is focused principally on the nucleocapsid Ag (Ferarri et al., 1990; Marinos et al., 1994; Chisari and Ferarri, reviewed 1995; Cao et al., 2001) and appears to be much more efficient than HBV envelope antigen with respect to activation of virus-specific T cells during acute HBV infection (Ferarri et al., 1990, Reignat et al., 2002, Hyodo et al., 2003). This is partly related to the fact that the humoral responses to the envelope Ags can show great variation from patient to patient during infection and after vaccination, ranging from high to non-responder phenotype (Milich et al., 1995a). Also, it is thought that T cell

responses to HBcAg could contribute in vivo to HBV clearance by amplifying and accelerating the development of neutralising anti-envelope antibodies (Ferarri et al., 1990). However, anti-HBc/HBe do not neutralize viral infectivity (Chisari and Ferarri, reviewed 1995).

1.7.3 CHRONIC HBV

For a non-cytopathic virus to persist, it must be able to evade immune surveillance; there must be either an ineffective antiviral immune response, or the virus must escape an otherwise efficient response. All of these might be involved in HBV persistency.

As tolerance against HBcAg/HBeAg eventually diminishes, there ensues a battle between virus and host, with different outcomes depending on immune status and HLA phenotype. Some progress has been made recently in teasing out the relationship between host and viral factors in this process. Viral persistence is generally thought to be due to an inadequate antiviral T lymphocyte response. This concept has derived from animal models of viral infection and from the study of patients who spontaneously clear the virus (in whom there is a strong virus-specific CD4 and CD8 response). This view is supported by several observations that patients with concurrent HIV infection or patients receiving immunosuppressive therapy are more likely to develop viral persistence (Heathcote et al., 1994).

1.7.3A HBeAg Seroconversion

Seroconversion from HBeAg to anti-HBe usually signifies a reduction in viral replication or clearance of the virus (Cariani et al., 1992; Carman et al., 1995a) and resolution of an acute case or the appearance of a milder, chronic infection with low or undetectable virus levels. The term HBeAg/anti-HBe seroconversion has become complicated by recent information that anti-HBe production can coexist with serum

HBeAg for long periods (years) in a very small proportions before the actual loss of HBeAg from the serum. It may be more accurate to view seroconversion as a long-term process in which HBeAg and anti-HBe coexist within immune complexes first in HBeAg excess and then, over time, in anti-HBe excess (Maruyama et al., 1998). In patients with chronic hepatitis during HBeAg/anti-HBe seroconversion there is evidence of a vigorous immune response characterised by $CD4^+$ T-helper-cell proliferation to HBcAg and emergence of anti-HBe (reviewed by Chisari and Ferrari, 1995). When HBeAg/anti-HBe seroconversion is followed by biochemical and histological remission, the subsequent low level of viral replication (detected only by PCR) implies that the dominant virus has been recognised and suppressed (or perhaps eliminated) by the immune response. The lack of serum HBV DNA in HBeAg positive patients can predict seroconversion to anti-HBe. However, when active disease persists with high level viremia, it is common that a virus (sometimes a pre-C mutant) has emerged that has evaded a critical element of the immune response (Carman et al., 1995a; Rodriguez-Frias et al., 1995; Maruyama et al., 1998).

1.7.3B HBV Variants

In populations worldwide, variants of HBV are selected after seroconversion to anti-HBe (Carman et al., 1989; Naoumov et al., 1992; Boner et al., 1995; Gerner et al., 1998). In some cases, years of anti-HBe positivity were necessary before selection of the pre-C stop codon variant (Carman et al., 1994). However, there is dispute regarding their degree of pathogenic potential. In some patients who showed persistence of serum HBV DNA despite seroconversion to anti-HBe, a translational stop codon at the 3' end of pre-C, predicting prevention of HBeAg synthesis, was reported in several studies on chronic HBV-infected patients (Carman et al., 1989; Naoumov et al., 1992; Maruyama et al., 1998). This variant has been found in

asymptomatic carriers (Akarca et al., 1994; Gandhe et al., 2003), in fulminant hepatitis (Hasegawa et al., 1994; Friedt et al., 1999), in patients with hepatocellular carcinoma (Takahashi et al., 1998; Zhong et al., 2000), after interferon (IFN) therapy (Gunther et al., 1992; Zhang et al., 1996) and after immunization (Lee et al., 1996). A_{1896} appears to be associated with severe liver disease in HBeAg negative patients with enhanced HBV replication (reviewed by Lok et al., 1997). However such a relationship between HBe-negative variants and severity of disease has not been confirmed by French and American researchers, who proposed epidemiological reasons for this correlation (Feray et al., 1993; Friedt et al., 1999). It is now recognised that the occurrence of A_{1896} is restricted to HBV genotypes with a T at nucleotide 1858 (reviewed by Lok et al., 1997), and this may account for its high prevalence in Asia and the Mediterranean basin, where the predominant HBV genotypes (B, C, and D) frequently have T_{1858} (Chan et al., 1999; Yoo et al., 2003). It has a low prevalence in the West, South and central America (Feray et al., 1993; Laskus et al., 1993), where the predominant HBV genotype A, F and H, almost always have C_{1858} (Inoue et al., 1998; Chan et al., 1999; Norder et al., 2003), which prohibits development of A_{1896} (Li et al., 1993; Inoue et al 1998; Sanchez-Tapias et al., 2002). An additional mutation A_{1899} accompanies A_{1896} occasionally (Carman et al., 1989; Laras et al., 1998). A_{1899} was also observed in isolates from several patients, but the functional implication of this mutation is still undefined (Cariani et al., 1992). A_{1899} is reported to be usually found with A_{1896} (Carman et al., 1989). Similarly, A_{1898} is linked with A_{1856} (reviewed by Carman, 1998). However, not all HBeAg negative carriers harbour the A_{1896} variant (Naoumov et al., 1992). It is proposed that the most common pre-C mutants (A_{1896}, A_{1899} and A_{1898}) enhance the stability of the lower stem of the encapsidation signal (ε) by replacing less stable T-G pairs between

nucleotides 1858-1896, 1855-1899 and 1856-1898, respectively, with more stable T-A pairs (Fig 1.2) (reviewed by Lok et al., 1997; reviewed by Papatheodoridis, 2001). Collectively, these authors suggested that failure to produce a target antigen may be a way to evade the clearance of infected hepatocytes, which in turn will explain viral persistence of the HBe minus mutants after seroconversion to anti-HBe (Carman et al., 1993). It seems likely that secretion of HBeAg represents a viral strategy to maintain a long-term infection without eliciting an overly destructive immune response, which would eliminate the virus and/or kill the infected host (Milich et al., 1997b). However, this is in contrast to the view that the severe liver damage is not due to the pre-C mutation per se, but to the high virus load despite the absence of the general marker for active infection, HBeAg, in serum (Hannoun et al., 2000a).

The serologic status after seroconversion of many chronic HBV patients is consistent with low level transcription of HBeAg-producing HBV DNA. Decreased levels of viral replication after seroconversion in these patients is marked by the inability to detect serum HBV DNA, decreased infectivity, and the decreased production of free HBeAg. It should be noted that the production of free HBeAg in chronic HBV patients can and does coexist with anti-HBe production. Therefore, it appears that HBeAg to anti-HBe seroconversion results from immune responses directed against HBeAg-expressing wild type virus, but this process does not necessarily involve the selection of e-minus HBV DNA variants. In fact, the low level of HBV infection that persists in many of these seroconverted-chronic patients is caused by pre-C wild type virus. This implies that small amounts of HBeAg co-exist with anti-HBe, possibly still modulating the immune response (Uy et al., 1986). This level of HBV replication may be below the threshold necessary to evoke an immune-mediated lytic response, and low level viral replication may be maintained by an ongoing antiviral immune

response (Maruyama et al., 1998). Alternatively, it may be sufficient to stimulate an immune response and consequently eliminate the wild type, prolonging virus persistence (Lee et al., 1996).

During chronic HBV infection, HBcAg is the only antigen that elicits a prominent immune response (Milich et al., 1997c), and in many reports, aa variability of HBcAg has been linked to the severity of liver disease, both in the HBeAg and anti-HBe phase, which leads to the hypothesis that patients who cannot clear the virus via a CTL response are susceptible to anti-HBc-mediated lysis of hepatocytes with the consequent selection of escape mutants (Carman et al., 1994, 1997b; Naoumov et al., 1997; Kratz et al., 1999). During the HBeAg-positive phase of chronic hepatitis, C-gene sequences have been reported to be relatively conserved, irrespective of disease activity (Carman et al., 1995a; reviewed by Lok, 1997). Selection of large numbers of aa substitutions within the core region seems to require not only seroconversion to anti-HBe, but also the previous selection of a pre-C stop codon.

Mid-core deletions have been found in patients with various types of chronic infection: in active liver disease (Wakita et al., 1991); in children with persistent viremia (Ni et al., 2000); in symptomatic carriers (Tsubota et al., 1998); and even in the HBeAg positive phase (Akarca et al., 1995b; Marinos et al., 1996). They have also been reported in HB-infected immunosuppressed patients (Gunther et al., 1996; Preikschat et al., 2002; Bock et al., 2003). Sequential analysis of core gene aa variability from the quiescent to the active phase of chronic infection showed that the mid core gene mutations may offer some advantage over viral replication and synthesis in persistent infection (Wakita et al., 1991; Tsubota et al., 1998; Bock et al., 2003). However, in some other studies these deletion mutants were reported not only to be replication defective, but also to suppress the replication of wild type (Marinos

et al., 1996; Ni et al., 2000). Core gene deletions were located within aa 81-113 (Ackrill et al., 1993); 79-112 (Akarca and Lok, 1995); 63-132 (Wakita et al., 1991); 20-132 (Ni et al., 2000), and 64-128 (Tsubota et al., 1998). These residues cover different epitopes: one CTL, two T helper, and three B cell (see section 1.6, table 1.4), and may change the immunological properties of HBcAg giving rise to immune escape.

Apart from pre-C and core variants, common changes during seroconversion include variants of the basal core promoter (BCP), which is embedded in the X gene. Deletions, insertions and point mutations (Chan et al., 1999) have all been described. Among them, T_{1762}/A_{1764} is common (Okamoto et al., 1994; Takahashi et al., 1998). Such mutations occur by substitution of A to T at nucleotide 1762, and G to A at nucleotide 1764, which may interfere with transcription of the HBeAg precursor (Okamoto et al., 1994; Laras et al., 2002). Although the effect of such mutations on the synthesis and secretion of HBeAg is not complete (Okamoto et al., 1997), Laras and colleagues found the absence or low levels of pre-C mRNA transcripts in patients who harboured these double mutations (Laras et al., 2002). Those mutations were significantly more common in patients infected with HBV who had C_{1858} (Chan et al., 1999). There is a striking presence of both BCP and pre-C mutants in anti-HBe positive infants with fulminant hepatitis (Friedt et al., 1999), however, the role of such an association is controversial because they have also been found in some HBeAg-positive patients without fulminant hepatitis (Carman et al., 1995a; Hou et al., 1999; Chu et al., 2003a). In Chinese symptomatic and asymptomatic chronic HB infection, it was suggested that T_{1762}/A_{1764} mutations were present or developed first to down regulate HBeAg expression, followed by introduction of a pre-C stop codon mutation, which then completely eliminated HBeAg synthesis (Hou et al., 1999). There is some

controversy in the literature concerning whether these mutations really enhance replication of HBV. As the pre-C gene and its products have also been reported to inhibit HBV progeny DNA synthesis (Laras et al., 1998), some hypothesised it is possible that one or several of the many substitutions interfered with the effects of the pre-C stop codon and the T_{1762}/A_{1764} mutations on viral replication (Sterneck et al., 1997, 1998).

Also, a number of X-region deletions and mutations have been described by a number of groups in chronic carriers (reviewed by Carman et al., 1998; Baumert et al., 1998; Baptista et al., 1999; Blackberg et al., 2003; Hwang et al., 2002, 2003). Finally, numerous missense and silent mutations in the preS1/S2 region were identified after anti-HBe seroconversion (Sanantonio et al., 1992; Gerner et al., 1998; Ding et al., 2003; Sugauchi et al., 2003), but their significance is unclear.

1.7.3C Chronic HBV and T-cell Response

In chronic carriers, the HBcAg specific Th response is significantly weaker, and in many patients, undetectable. This difference in the HBcAg-specific Th activity between acute and chronic HBV infection supports the concept that the Th cell response to HBcAg may influence the outcome of HBV infection (Hyodo et al., 2003). HBcAg-specific Th cell non-responsiveness (or tolerance) is reversible, enhancing both the humoral and cellular effector immune response to HBV in chronic patients. The induction of this response may play a role in the transition from the immunotolerant phase of chronic HBV to the active phase, and eventually seroconversion (Marinos et al., 1994).

Milich and co-workers (1986) observed that the two forms of the HBV nucleoprotein preferentially elicited different Th cell subsets and may have profound implications in terms of the mechanisms of viral persistence and immunopathogenesis. Many

chronically infected patients maintain IgM anti-HBc, although usually at low titres. The switch from predominantly IgM to high titres of IgG anti-HBc requires T-cell help function, which may be variably present from patient to patient and defective in patients who progress to chronicity, as the T cell response to HBc/HBe antigens is strong during acute hepatitis and weak in chronically infected patients. Thus, a strong antibody response in chronically infected patients may be due to the known T cell independence of HBcAg. This would explain the slow decline in IgM anti-HBc titres during the first 1 to 2 years of chronic infection. Since T cell recognition of HBcAg and HBeAg is highly cross reactive, T-cell help for IgG anti-HBc production would presumably also stimulate anti-HBe production. One would then predict that the decline of IgM in favour of IgG anti-HBc should correlate with anti-HBe seroconversion (Milich et al., 1986). HBcAg elicited primarily IgG2a and IgG2b anti-HBc antibodies, whereas the anti-HBe antibody response was dominated by the IgG1 isotype (Milich et al., 1997c). These results indicated that HBcAg preferentially, but not exclusively, elicits Th_1-like cells and that HBeAg preferentially, but not exclusively, elicits Th_0-or Th_2-like cells. These results may have relevance in chronic HBV infection since circulating HBeAg may downregulate antiviral clearance mechanisms by virtue of eliciting anti-inflammatory Th_2-like cytokine production (Milich et al., 1997a).

Other groups have shown that host genetic background determines Th_1 v Th_2 switching. Some alleles, such as DR3, DR13 and DR14 preferentially induce a Th_1 (cellular) response, which would inhibit the Th_2 (humoral) response and limit the production of anti-HBs antibodies (Caillat-Zucman et al., 1998; Cao et al., 2001). Conversely, the response would be skewed towards Th_2 responsiveness in subjects with DR1 and DR2 alleles (Caillat-Zucman et al., 1998).

CTL exhaustion (or anergy) has been proposed as being involved in HB persistence. Chisari and Ferarri, (1995) have shown that the number of potentially infected hepatocytes in the human liver (~$1X10^{11}$), compared to the number of CD8 (~25% of total body CTL population of ~$1X10^{12}$) would require an extraordinary high proportion of the total body CTL to be in the liver. This exhaustion of antiviral CTL stimulated by high dose antigen is supposed to occur in HB infection acquired in adulthood, whereas clonal deletion of HBV-specific T cells as a consequence of transplacental HBV infection is believed to be involved in vertically infected patients. Both of these factors reduce the number of HBV-specific T cells in chronically infected patients. However, in a recent study using MHC/peptide tetramers and ICCS (intracellular cytokine staining) methods, Reignat et al demonstrated that a population of envelope-specific CD8 cells escaped exhaustion mediated by high concentration of antigen. In addition, they showed that core (aa 18-27) and envelope (aa 183-191)-specific CD8 cells were able to expand, displaying the same proliferation potential, activation threshold and IFN-gamma production ability as memory CD8 cells present in immune individuals. These cells were characterized by altered HLA/peptide tetramer reactivity and could be reversed with repetitive stimulation. They suggested therefore that these cells were not "anergic", but, "ignorant" instead (Reignat et al., 2002).

T cell exhaustion and mutational escape have been shown to be involved in other persistent virus infections such as HCV, HIV and LCMV. Escape through "speed" has been demonstrated as a strategy of RNA viruses leading to high level antigen persistence in the presence of low levels of CD8 and CD4 T cells in human and animal models (Gallimore et al., 1998; Lenchner et al., 2000; Ou et al., 2001, reviewed by Lucas, 2001). CTL-escape mutations have also been described

extensively in these viruses (Lechner et al., 2000; reviewed by Sewell, 2000; reviewed by Lucas, 2001; reviewed by Klenerman, 2002) although with some differences; e.g., only infection with high doses of LCMV results in the generation of CTL escape variants (Pircher et al., 1990).

1.7.3D Chronic HBV and HBcAg Localisation

For liver cells to become a "target cell" at least theoretically there should be viral antigens in the cell membranes before the process of necrosis. The significance of the membranous expression of viral antigens in the pathogenesis of hepatic injury has been studied extensively and some authors suggested a strong correlation between membranous expression of HBsAg with overall activity of liver disease (Ray et al., 1976; Gowan et al., 1983; Chu et al., 1988; Chu et al., 1995 a and b; Sharma et al., 2002). The same finding has been proposed for membranous expression of HBcAg during chronic HB infection (Wu et al., 1993), as the blocking effect of anti-HBc was attributed to the binding of the antibody to the corresponding antigen on the surface of the infected hepatocytes (Mondelli et al., 1982). Variable patterns of HBcAg expression in different forms of liver damage have led investigators to postulate that this expression may be a prerequisite for immune mediated liver cell damage (Ray et al., 1976; Chu et al., 1987; Wu et al., 1996). Differential expression of the HBcAg and HBsAg at the cellular level and the ultimate development of different histological types of hepatitis reported so far (Ray et al., 1976; Huang et al., 1979; Gowan et al., 1983; Chu et al., 1994; ter Borg et al., 2000; Sharma et al., 2002), has led to the hypothesis of a combination of humoral and cellular immune response in the pathogenesis of liver cell injury.

The localisation of HBcAg has significant prognostic implications, as follows: first, core Ag possesses unique immunological features (Milich et al., 1997a), and C-gene

products are major targets of the host immune response, particularly in chronic infection (Mondelli et al., 1982; Chu et al., 1987; Ferarri et al., 1987, 1992; Milich et al., 1989). Second, the pattern of the intracellular distribution of C-gene as detected by immunostaining, has been associated with different levels of virus replication, liver disease activity, and liver cell regeneration (Ray et al., 1976; Akiba et al., 1987; Wu et al., 1993; Ballare et al., 1989; Tong et al., 1990; Chu et al., 1994, 1997; Bock et al., 2003). Third, the intracellular accumulation of non-enveloped core particles can be cytopathic (Roingeard et al., 1990a). Fourth, the presence of mutant core protein in liver cells can interfere with the efficiency of virus production (Scaglioni et al., 1994; Bock et al., 2003), and possibly with response to IFN therapy (Naoumov et al., 1995). The localization of HBcAg in hepatocytes varies even in the same tissue (Akiba et al., 1987). Many reports in the literature suggested that HBcAg may be expressed in the cytoplasm, nucleus or both in infected or transfected cells (Yamada et al., 1977; Brechot et al., 1980; Gowan et al., 1983; Uy et al., 1986; Wu et al., 1987; Akiba et al 1987; Hsu et al., 1987; Fang et al., 1994; Bock et al. 1996; Chu et al., 1997; Naoumov et al., 1997; Park et al., 1999; ter Borg et al., 2000; Sharma et al., 2002; Bock et al., 2003). In fact, although early reports had shown that HBcAg was located in the nucleus (Ray et al., 1976; Huang et al., 1979), later reports have made clear that HBcAg can be seen in the nucleus or cytoplasm, or both.

Various markers have been postulated to identify patients with a higher general level of HBV replication during chronic HB: the presence of detectable HBV DNA (or HBV virion) in serum, and expression of HBcAg within the cytoplasm of infected hepatocytes (Brechot et al., 1980, 1981; Gowan et al., 1983; Chu et al., 1987; Hsu et al., 1987; Ballare et al., 1989; Ou et al., 1990; Chu et al., 1995). In different patients, the above markers of virus replication tend to co-exist, and such patients also tend to

be those with histological evidence of active disease. For example, topographic distribution of HBcAg was closely related to HBeAg/anti-HBe status (Ballare et al., 1989; Ou et al., 1990). Liver cell HBcAg expression can be correlated with the three phases of chronic HBV infection: first, HBeAg positivity, immune tolerant phase, characterised by nuclear HBcAg, mild disease and high viral replication; second, HBeAg seroconversion, immune clearance phase, with virus replication/elimination, cytoplasmic and/or membranous HBcAg, and third, inactive virus replication phase with negative HBcAg (integrated HBV DNA into host cell), and a bipolar disease spectrum (either normal histology, or, cirrhosis and HCC) (Hsu et al., 1987; Wu et al., 1993; Chu et al., 1997; Serinoz et al., 2003). However, in some patients with circulating anti-HBe, no detectable circulating HBV DNA and less active forms of liver disease have been reported: intracellular HBV DNA tends to be in a high molecular weight form reflecting integration into host cell DNA (Brechot et al., 1981; Bonino et al., 1981; Hadziyannis et al., 1983; Gowan et al., 1983; Wu et al., 1987, 1996). The state of HBV DNA in the hepatocytes (free or integrated into the host genomes) (Brechot et al., 1981), and the presence of the HBcAg in liver tissue represent active replication of HBV. This concept was based on the observations that a high concentration of serum HBV DNA has been found in patients with positive HBcAg-staining hepatocytes (Chu et al., 1994; Wu et al., 1996). Further, Gowan et al. found a huge amount of HBV DNA, much of it in a single-stranded form, in the cytoplasm of infected hepatocytes in active cirrhosis. They suggested that the large pool of non-encapsidated cytoplasmic HBV DNA indicated cells undergoing replication with or without mature virion formation. Then they found a relation between HBcAg expression and this form of DNA; cytoplasmic HBcAg was highly predictive of the presence of large amounts of cytoplasmic DNA in the same cells,

whereas, nuclear HBcAg was seen in cells with and without such levels of DNA, presumably undergoing latent or abortive infection (Gowan et al., 1983).

Although some authors believed that expression of HBcAg in the hepatocyte nucleus is helpful for estimating viral replication in chronic HBV infection (Wu et al., 1996; Chu et al., 1997; Lindh et al., 1999; Serinoz et al., 2003), many others suggested that cytoplasmic HBcAg was more appropriate as a marker of disease activity (Hsu et al., 1987; Kakumu et al., 1989; ter Borg et al., 2000). Therefore, the pattern of preferential cytoplasmic core staining in different studies has led investigators to several speculations. The cytoplasmic expression of cAg was suggested to be due to the presence of a particular virus population characterised by multiple mutations in the C-gene, which possibly inactivated the function of the nuclear localisation signals of the core protein (see below). Mutations may also modulate envelope protein interaction with the core protein (Yeh et al., 1994). Point mutations of the pre-C region is suggested to prevent nuclear transport of HBcAg. This pattern has been shown by a study by Park and colleagues on chronic HB patients who had the A1896 mutation; they found cytoplasmic expression without a single example of nuclear expression (Park et al., 1999).

Another mechanism for cytoplasmic localisation of HBcAg might be related to cell cycle regulation. In the G0/G1 phase, HBcAg is predominantly localised in the nucleus, whereas, in the S phase, the amount of HBcAg in the nucleus is greatly reduced and HBcAg found almost entirely in the cytoplasm (Yeh et al., 1993). This hypothesis suggests that the differential subcellular localisation of HBcAg in chronic HBV infection is related to the activity of hepatocyte regeneration. Chu and co-workers demonstrated that almost all of the hepatocytes with nuclear expression of HBcAg were quiescent and resting, whilst most of the hepatocytes with cytoplasmic

expression of HBcAg were regenerating and proliferating (Akiba et al., 1987; Wu et al., 1987, 1993; Ballare et al., 1989; Tong et al., 1990; Chu et al., 1995, 1997; Serinoz et al., 2003). On the contrary, Naoumov et al (1997) speculated that the influence of the cell cycle observed in cultured cells appeared to play only a minor role in vivo. Even in patients with cirrhosis and severe chronic hepatitis, with hepatocytes in different stages of the cell cycle because of liver cell regeneration, core gene products predominantly had nuclear localisation, similar to the finding in patients with minimal inflammatory reaction in the liver.

Taken together, the correlation between the degree of expression of HBcAg in the nucleus and/or cytoplasm and the level of liver disease activity is relative, not absolute. This conclusion is based on some observations on patients who showed a high degree of liver disease activity in the presence of nuclear HBcAg expression, and on the other hand, mild hepatitis in the presence of cytoplasmic HBcAg expression (Hadziyannis et al., 1983; Akiba et al., 1987; Chu et al., 1997).

There is still controversy about the correlation between HBcAg localisation and immune hepatocytolysis, as some investigators speculate that a shift of HBcAg from nuclear to cytoplasmic is secondary to immune-mediated liver cell injury rather than triggering the immune system by itself (Akiba et al., 1987; Chu et al., 1995). However, because cytoplasmic/membranous expression of HBcAg correlates rather well with the high degree of liver inflammatory activity (Akiba et al., 1987; Wu et al., 1987, 1993, 1996; Ballare et al., 1989; Tong et al., 1990; Chu et al., 1995), it is therefore suggested that hepatocytes with cytoplasmic/membranous expression of HBcAg represent target cells for immune hepatocytolysis, which would be followed by both subsequent seroconversion from HBeAg to anti-HBe and disappearance of

HBcAg and regeneration of hepatocytes (Ray et al., 1976; Chu et al., 1987; Wu et al, 1993; Park et al., 1999).

1.8 HBV EVOLUTION

The origin(s) of viruses cannot be known with certainty. PCR and other sensitive molecular techniques will reveal some viral genome sequences from the relatively recent past, but very ancient viral genomes will remain a matter for speculation. Comparative sequence analysis suggests that both RNA and DNA viruses have deep, archaic evolutionary roots both for genome structural organisation and as regards certain genomic and protein domains. Both DNA and RNA viruses can emerge and evolve by a variety of mechanisms including mutation, recombination and reassortment (Holland et al., 1998). Nucleotide substitutions in viral genomes can have several effects, including evasion of vaccine-induced or natural immunity, drug resistance, and changes in pathogenicity, alteration in tissue or species tropism, and viral persistence.

Escape mutation is the process by which aa substitution at one or more positions alters the epitope to the extent that the virus can persist in the presence of an adequate immune response to the initial epitope. As recognition of the first epitope leads to the destruction of infected cells or neutralization of free virus, viruses that encode for the expression of a mutated epitope (or cells that present it on their surface) will survive. Viruses that are replicating more rapidly are more likely to develop mutations and for them to be selected (reviewed by Carman et al., 1995c). A comparison of the rates of synonymous and nonsynonymous substitution in viral genomes can elucidate, to some extent, whether there has been selection at the aa level. One would expect the rate of synonymous to be less than the rate of nonsynonymous substitution if there is selective pressure exerted on this viral gene (Mizokami and Orito, 1999).

1.8.1 HBV Mutation Rate

Although the reverse transcriptase activity of HBV may be responsible for the high rate of nucleotide substitution of the virus compared to that of other DNA viruses, replication of the HBV genome may not always depend upon reverse transcription and thus, the frequency of mutation of the HBV genome during replication may not be as high as that of a retroviral genome (Kidd-Ljunggren et al., 1999). This is due to the fact that the mutation rate of HBV is influenced by some related factors. Firstly, the genome is very compact with overlapping ORFs, which limits the number of viable mutations (Morozov et al., 2000). In a virus with overlapping open reading frames, many mutations will generate non-viable virions though some mutations may give the virus a replication advantage (reviewed by Torresi, 2002). The second constraint is imposed by the need to conserve the direct repeats, promoters and other cis acting elements, which are involved in replication of the genome. Thirdly, HBV replication takes place inside the nucleocapsid and, unlike retroviruses, this process involves only one copy of the RNA pregenome, limiting the chance of homologous recombination (Morozov et al., 2000, Sugauchi et al., 2002). Despite these constraints, up to 12% of nucleotides may vary between isolates of HBV (reviewed by Carman et al., 1997c). The net result will be a higher mutation rate in unconstrained parts of the genome. One such region is the pre-S, especially pre-S2 (Kidd-Ljunggren et al., 1999). Considering the rate of synonymous substitution is higher than nonsynonymous substitution for all HBV-ORFs (Yang et al., 1995), it has been suggested that HBV genome variation is constrained by aa changes (Orito et al., 1989). This led to the hypothesis that each protein or each domain of an ORF has its own functional and structural constraints which conserve the aa sequence over evolution time. The assumption is that the selective force on each ORF works

independently of the region of gene overlap and that base exchanges in the overlapping regions are therefore the result of dual selection by the two overlapping ORFs (Kodama et al., 1985). The appearance of mutations in overlapping but unrelated viral genes (for example S and P genes) may produce HBV mutants with altered antigenicity and/or replication and a natural history that may be distinctly different to wild type HBV (reviewed by Carman et al., 1997; reviewed by Torresi, 2002).

The rate of replication and mutation is an important factor in the genetic evolution of a virus. Some DNA viruses, such as hepadnaviruses, which include a reverse transcription step in their replication cycle, show rates of evolution in the range of 10^{-4} to 10^{-5} substitution per nucleotide per year, which is approaching the values of some RNA viruses. HBV exhibits a mutation rate more than 10-fold higher than other DNA viruses. The estimated rate of evolution for HBV is $<2x10^{-4}$ base substitutions/site/year, which reflects a highly dynamic process with a large production of virus (reviewed by Domingo et al., 2001). This value is intermediate between DNA and RNA viruses (Morozov et al., 2000). Although the rate of synonymous substitutions for HBV is 10^4 times higher than that of a host genome, it is 10^{-2} less than that of retroviral genes (Orito etal., 1989). Nevertheless, this rate provides support for the hypothesis that HBV evolved from a retrovirus or retrovirus-like progenitor through a process of deletion (Miller and Robinson, 1986; Okamato et al., 1987).

The HBV genome is extremely stable unless exposed to host immune responses, as exemplified by the completely conserved nucleotide sequence over a 20-35 year period in HBeAg-positive asymptomatic carriers with very high levels of virus replication. In contrast, mutations are seen in most, if not all, HBeAg-negative carriers

and are distributed over all regions of the viral genome (Hannoun et al., 2000a). Some of these aa substitutions can be found as minor species at earlier time points suggesting that the selection process is slow and that multiple strains can co-exist (Carman et al., 1996). Some observations have also shown that if conditions do favour the emergence of a variant (for example in the case of a poor match vaccine between immunity and the virus), then the time-scale of emergence is likely to be in the order of decades. As HBV persists in individuals infected for long periods spreading the infection relatively slowly, a variant will only emerge slowly (Wilson et al., 1999).

1.8.2 Selection of the fittest strain

The term "fixation" refers to random mutations not being lost but becoming incorporated into sequences that undergo further rounds of replication and become progeny virions (Carman et al., 1993). There are two ways by which fixation occurs: one is the occurrence of mutations that do not affect fitness, which has been shown in cell culture in which viral populations gradually shift their consensus sequence while exposed to a constant in vitro environment. Lenhoff et al (1998) generated a cytopathic mutant of DHBV (G133E) in the pre-S protein of DHBV. Inclusion of this mutant into susceptible ducklings resulted in enhanced viral replication, increased the pool of viral cccDNA, and caused hepatocyte destruction. The liver damage caused by G133E DHBV subsided over time resulting in mild chronic hepatitis similar to that observed in wild type virus-infected birds and coincided with a reduction of viral replication to wild type virus levels in the liver. Further, they identified at least one noncytopathic revertant from the serum of G133E-infected birds after recovery, suggesting that acute liver injury could result from infection with a cytopathic hepadnavirus, but such viruses may be rapidly replaced by noncytopathic variants during persistent infection. In other words, these cytopathic viruses are not as

replication fit as the wild type in the context of persistent infection (Bartholomeusz and Locarnini, 2001).

The alternative mode of mutant fixation is positive selection of fitter virus as driven by antibodies, antiviral agents, or differences in cell biology. In the natural course of HBV infection, cellular and humoral immune pressure against virus-specific proteins may select the fittest viral strains (Liu et al., 2002). The type and number of mutations that accumulate in an individual genome are either a marker of the duration and/or severity of the liver disease, or the type and the intensity of immune response (Gunther et al., 1999). Selection by the host is a major force for evolutionary change within a virus population (reviewed by Carman et al., 1998). Changes in host selection pressure may greatly affect substitution rates in HBV, with lower rates of changes in those individuals who continue to produce the viral HBeAg compared to those who have cleared it (Carman et al., 1995a; Okamuru et al., 1996; Bozkaya et al., 1996, 1997). In numerous studies, investigators have shown that during the HBeAg-positive active replication phase of chronic infection, despite the virus being capable of mutating because of the poor proof-reading ability of reverse transcriptase, sequences able to translate HBeAg remain dominant. Possibly, HBeAg-producing sequences have an intracellular advantage over HBeAg-nonproducers. Because of the dominant replication efficiency of HBeAg-producing strains and thus tolerance, few cells appeared to be destroyed before the elimination phase, therefore, there appears to be no selection pressure and thus e-minus mutants are lost or remain as minor populations. However, during the elimination phase of chronic disease, there is the added factor of positive selection pressure, and hepatocytes with HBeAg-nonproducing strains become dominant. At the molecular level, this is explained by the presence of pre-C mutants (Carman et al., 1996b). Moreover, it has been shown

that once the resistant virus became predominant in the viral quasispecies obtained after treatment with lamivudine, drug removal led to the rapid replacement of the resistant virus by the wildtype (Abdolhamed et al., 2002). Other studies showed that the pre-C and core promoter mutations were replaced by wild type during long term therapy, but with continuing therapy, mutations reappeared independent of viral breakthrough (Cho et al., 2000; Suzuki et al., 2002). Selection is likely to occur if the immune response is incomplete (by using vaccines or monoclonal antibody immunotherapy) (reviewed by Domingo et al., 2001).

1.8.3 HBV Origin

Attempts have been made to relate these observations to HBV evolution. Much of our knowledge about human hepatitis has relied heavily upon information derived from infection (natural and experimental) of non-human primates during the last 50 years. However, there seem to be some difficulties due to the relationship between the human genotypes A-E and G (and even more divergent genotype F) to each other and to other primate-associated genotypes. The origin of HBV in humans is as confusing as that of the hepadnaviruses from other primates. Different theories have been proposed by investigators on HBV origins based on the hypothesis that the numbers of nucleotide and aa substitutions over time, the molecular clocks, are indicators of viral evolution (Mizokami et al., 1999). A primate origin for HBV infection was proposed by MacDonald et al., 2000. This theory based on the finding of variants in chimpanzees (McDonald et al., 2000), woolly monkeys (Lanford et al., 1998), and orangutans (Warren et al., 1999), suggested that these viruses co-evolved with their primate hosts over periods of 10-35 million years. This hypothesis has been supported by the observations that areas of high HBV prevalence in humans are those in which contact with, and cross-species transmission from primates are most likely (South

America, Sub Saharan Africa and Southeast Asia). Indeed, certain HBV genotypes are specific to these three areas (F, E, B/C, respectively). Moreover, the mixture of HBV genotypes found outside these areas, such as Europe and North America, may result from much more recent epidemic spread (Simmonds et al., 2001a). In contrast, a recent emergence hypothesis for HBV infection indicated that the current wide distribution of HBV in apes must have arisen through several cross-species or subspecies transmission in the relatively recent past (Starkman et al., 2003)

Based on Norder and her co-worker's observation that most of the dendrograms obtained from gibbon and chimpanzee strains represented early lineages, assumptions were made that these viruses were indigenous to their respective hosts and not recent acquisitions from man (genotype F). Therefore, they suggested that either genotype F represents an early cross-species transfer from a non-ape primate to man, or that an hepadnavirus of a common ancestor to man and apes gave rise to two viral lineages (Norder et al., 1996). Thus, they proposed that the evolutionary history of HBV corresponds to the spread of anatomically modern humans as they migrated from Africa 100,000 years ago (Norder et al., 1994; Magnius and Norder, 1995), and different genotypes infecting humans evolved since this dispersal. However, this hypothesis does not explain the origin of the various non-human primate viruses which are interspersed among the human genotypes in the phylogenetic tree. The phylogenetic tree of the various primate HBV variants in no way reflects the phylogeny of the host species, as would be expected for co-speciation (Simmonds et al., 2001b). For example the presence of genotype F in native American populations is inconsistent with the presence of genotype B and C in their genetically nearest relatives, the Mongoloid Northeast Asian. Indeed, there is little relationship between HBV genotype distribution with any of the other human population groups (Southeast

Asian, Caucasians, and African population) (Simmonds et al., 2001b). Alternatively, the HBV genotypes may have evolved later than, and independent of, human migration (Gunther et al., 1999).

According to the finding that HBV showed a substitution rate of $2.1x10^{-5}$ substitutions per site per year over a mean observation period of 22 years, Orito and co-workers proposed that the human genotypes of HBV would have originated from a common ancestor approximately 3000 years ago (Orito et al., 1989). In this study, they showed that three major clusters of HBV (birds, mammals and humans) diverged from their common origin in the same order as that of host evolution. They concluded that the evolution of the hepadnavirus family was independent of host-species divergence and for HBV in humans this has taken place much more recently than has divergence of humans.

Alternatively, a New World origin for HBV was proposed by Bollyky and colleagues who suggested that HBV originated from the Americas and spread into the Old World over the last 400 years after contact from Europeans during colonisation; a genotype F origin. Further, they considered the possibility that if the virus originally entered the Americas from Asia, this may have required a higher rate of nucleotide substitution as it adapted to this naïve human population (Bollyky et al., 1999). However, the main problem for this hypothesis is the observation of the widespread distribution of HBV in Old World primate species. A remarkable example is a shared genotype of HBV infecting West African chimpanzees (McDonald et al., 2000; Takahashi et al 2000), which showed approximately 11% divergence from the human genotypes A-E. This finding was based on analysis of mutations in the C-terminus region of the core protein (which is well conserved among hepadnaviruses) between human genotypes E/F and the chimpanzee one. Interestingly, HBV-E/F and the non-human primate

hepadnavirus had a common motif within 20 nucleotides upstream from the stop codon for the core gene, whereas, HBV-A/B/C/D genotypes contained a different motif at this site (Takahashi et al., 2000). It has been revealed that sequences in wild-born Old World primates from Africa and Southeast Asia were unrelated to five human HBV strains (A to E) ; the conclusion is that the virus was not acquired from humans, and all the Old World non-human primate HBVs were on a common ancestral branch (Robertson et al., 2001). This finding together with the observation that the closest relative of the woolly monkey HBV is genotype F (Lanford et al., 1998), led to the speculation that chimpanzees have their own hepadnavirus, which resembles the human hepadnavirus (genotype F) (Takahashi et al., 2000; Robertson et al., 2001).

At present, the problems associated with each of these hypotheses for the origin of HBV prevent a definitive conclusion. Resolution of these issues requires more extensive sequence analysis of HBV in poorly sampled areas as well as combined human and primate studies together with utilizing models of DNA substitution which better describe the process of viral evolution (Bollyky et al., 1999).

1.8.4 HBV-genotypes evolution

The genetic variability of HBV is observed both as the evolution of genotypes (and thus subtypes), i.e. a divergence of the viral genome in the carrier population, and as the emergence of mutations in each infected subject (Lindh et al., 1999). Subtypes of HBV have largely evolved separately with their hosts over time and random mixing of different subtypes within individuals has been rare (Kodama et al., 1985). Okamoto et al (1987) observed that a 54-year old patient was chronically infected with three

different clones of HBV. They suggested that the clones would have evolved from a common ancestor virus during 54 years after the primary infection rather than infection by three different strains of HBV (triple infection). Further, they suggested that following a small-dose infection, only one genotype of virus, predominant in the donor, would most likely infect the recipient. In another study the predominant HBV genotype was shown to be quite stable over a period of some 30 years (Blackberg et al., 2000).

Chapter 2 Materials and Methods

2.1 Materials

2.1.1) DNA Extraction from serum by High Pure Viral Nucleic Acid Kit (Roche Diagnostics, Lewes, East Sussex, UK)

Binding Buffer	6M guanidine HCl, 10mM Urea, 10mMTris-HCl, 20% Triton® X-100 (v/v), pH 4.4
Poly (A) carrier RNA	0.2mg/40ul (after reconstitution)
Proteinase K	20mg/ml (after reconstitution)
Wash buffer	20mM NaCl, 2mMTris-HCl (pH 7.5)
Elution buffer	Nuclease-free redistilled H_20
High Pure filter tubes	Polypropylene tubes have two layers of glass fiber fleece and can hold up to 700ul of sample volume
Collection tubes	2ml Polypropylene tubes

2.1.2) Synthetic oligonucleotides

Oligonucleotides used during this work are listed in tables 2.1. Specific primers were used for amplifying C and S genes (conventional PCR) (table 2.1A). For mutagenesis study, oligonucleotides designed in length between 24 to 39 bases with 5' end phosphorylated (table 2.1B). The mismatch portion was at or near the 5' end of each primer with more than 20 correct bases on the 3' end (purchased from MWG, Biotech, Milton keynes).

2.1.3) Enzymes

Taq DNA polymerase, restriction enzymes and T4 DNA ligase were obtained from Roche Diagnostics. RNase was purchased from Sigma. EcoR1 and Hind III from Life Technologies.

Table 2.1 (A) Primers used in this study.

Primer	Gene	Sequence 5'⟶ 3' of Oligonucleotides	Base Position	Type
C1	Core	CGG GAT CCG AGG AGT TGG GGG AGG AGA TT	1726-1754	Sense
C3A	Core	G(R)T CTR TGT A(W)T AGG AGG CTG (R=A/G) (W=A/T)	1763-1783	Sense
C4	Core	CCT TAT GAG TCC AAG G(R)A TA (R=A/G)	2478-2459	Anti-sense
S1	Surface	CCT GCT GGT GGC TCC AGT TC	56-75	Sense
S2Na	Surface	CCA CAA TTC (K)TT GAC ATA CTT TCC A (K=G/T)	1003-979	Anti-sense
S6C	Surface	GCA CAC GGA ATT CCG AGG ACT GGG GAC CCT G	113-146	Sense
S7D	Surface	GAC ACC AAG CTT GGT TAG GGT TTA AAT GTA TAC C	857-823	Anti-sense

Oligonucleotide primers used for PCR and sequencing. Base positions numbered from the EcoRI site.

Table 2.1 (B) Oligonucleotide primers designed for ExSite™ PCR-Based Site-Directed Mutagenesis.

Primer Code	Sequence 5'⟶3' of oligonucleotide	Core Gene Base position	Type
50-B1.F	GGT GTT AAT TTG GAA GAT CCA GCA TCT	2117-2144	Sense
50-B1.R	CAC CCA GGT AGC TAG AGT CAT TAG	2116-2092	Anti-sense
50-B2.F	CCA GCT TAT AGA CCA CCA AAT GCC CCT	2288-2315	Sense
50-B2.R	AGG AGT GCG AAT CCA CAC TCC GAA AG	2287-2261	Anti-sense
I63-B.F	TTG GAA GAT CCA GCA TCT AGA GAC CTA GTA GTC AG	2126-2161	Sense
I63-B.R	ATT ACC ACC CAC CCA GGT AGC TAG	2125-2101	Anti-sense
I63-C.F	AGG TCC CCT AGA AGA AGA ACT CCC TCG CCT	2360-2390	Sense
I63-C.R	GCC TCG TCG TCT AAC AAC AGT AGT CTC	2359-2332	Anti-sense
92-B.F	CCA GCA TCT AGG GAC CTA GTA GTC AGT TAT GTT AAC ACT	2135-2174	Sense
92-B.R	ATC TTC CAA ATT AAC ACC CAC CCA GGT AGC	2134-2104	Anti-sense
I128-B.F	GGC AAT AAT TTG GAA GAT CCA GCA TCT AGG	2115-2145	Sense
I128-B.R	CAC CCA GGT AGC TAG AGT CAT TAG GTC	2114-1087	Anti-sense
I128-C.F	GAA TCT CAA TGT TAG TAT CCC TTG GAC	2437-2464	Sense

I128-C.R	CCG AGA TTG AGA TCT TCT GCG ACG CGG	2436-2409	Anti-sense
I131-B2.F	GTA **TCT** TTC GGA GTG TGG ATT CGC ACT CCT	2258-2288	Sense
I131-B2.R	CAA ATA CTC TAT AAC TGT TTC TCT TCC	2257-2230	Anti-sense
I131-C.F	AGA **A**CT CCC TCG CCT CGC AGA CGA AGG TCT	2374-2404	Sense
I131-C.R	TCT TCT AGG GGA CCT GCC TCG TCG TCT AAC	2373-2343	Anti-sense
I138-C.F	TCT **C**AA TCT CGG GAA TCT CAA TGT	2428-2452	Sense
I138-C.R	TCT TCT GCG ACG CGG CGA TTG AGA CCT	2427-2400	Anti-sense

Each primer code represents: sample code as well as primer direction; F, forward; R, reverse. Numbers indicate: 1 and 2 represent B cell epitope residues 74-89 and 128-135, respectively. Mutated bases shown by bold letters. Base positions numbered from the EcoRI site.

2.1.4) Reagents and buffers for PCR

10x PCR buffer	200mM Tris-HCl (pH 8.4), 500 mM KCl
10x dNTPs	100mM of each dATP, dCTP, dGTP, dTTP
TaqStart Antibody	1.1 ug/ul in storage buffer: 50mM KCl, 10 mM Tris-HCl (pH 7.0), 50% glycerol
TaqStart dilution buffer	50mM KCl, 10mM Tris-HCl (pH 7.0)
10x TBE	89mM Tris HCl (pH 8.0), 89mM boric acid, 1mM EDTA (Ethylenediaminetetra- acetic acid (disodium salt)
10x agarose gel loading buffer	1x TBE, 1% SDS (Sodium dodecyl sulfate), 50% sucrose, 1 mg/ml bromophenol blue
X-gal (10%)	5-Bromo-4-chloro-3-indoyl-β-D-galactopyranoside
IPTG 100 mM	Isopropyl-1-thio-β-D- galactopyranoside

2.1.5) Materials provided in ExSite™ PCR-Based Site-Directed Mutagenesis Kit

ExSite™ DNA polymearse blend (5 U/ul)

Cloned pfu DNA polymearse (2.5 U/ul)

Dpn I restriction enzyme (10 U/ul)

T4 DNA ligase (4U/ul)

10 x mutagenesis buffer

10 mM rATP

25 mM dNTP mix (6.25 mM each dNTP)

pWhitescript™ 4.5-kb control template (1.5 ug/ul)

Oligonucleotide control primer #1[phosphorylated 24-mer (75 ng/ul)]

Oligonucleotide control primer #2 [30-mer (75 ng/ul)]

Epicurian Coli® XL-Blue supercompetent cells

pUC 18 transformation control plasmid (0.1 ng/ul in TE buffer)

2.1.6) Bacterial strain and growth media

Escherichia coli. Genotype:Φ80dlacZΔM

DH5α	15, rec A1, endA1, gyrA96, thi-1, hsdR17 (r_k^-, m_k^-), supE44, rel A1, deoR, (lacZYA-argF) U169
L-broth	Luria-Bertani liquid medium (10 g NaCl, 10 g Bacto Tryptone, 5 gm Bacto-yeast extract in one litre distilled water.
L-broth in agar	As above plus 10 g Bacto-agar

2.1.7) Plasmid

pRK5 was used to express full-length core gene in COS-7 mammalian cells. PRK5 contains: a cytomegalovirus (CMV) promoter which allows very high expression of proteins in a variety of mammalian cells, a multiple cloning site (MCS) downstream of the promoter, a SV40 origin of replication for single stranded DNA, and an ampicillin resistant gene (AMP^R) for prokaryotic selection (a kind gift from professor H. Will, Hamburg, Germany).

2.1.8) COS7 culture system

These monkey kidney cells were originally derived from an African –monkey kidney cell line transformed by an origin defective mutant of SV-40. Cells were grown in Dulbecco's

modified Eagle medium (DMEM) supplemented with 10% bovine calf serum, 100 IU/ml penicillin/ 100ug/ml streptomycin and 2mM glutamine. Trypsin (0.25% trypsin dissolved in Tris-saline) and Versene (600mM EDTA in PBS A, 0.0015% (w/v) phenol red) were used for splitting of the confluent cells.

2.1.9) Reagents and solutions for small scale plasmid preparation

Miniprep Solution I	25mM tris HCl (pH 8.0), 50 mM glucose, 10mM EDTA (pH 8.0)
Miniprep Solution II	0.2M NaOH, 1% SDS
Miniprep Solution III	3M K. acetate, 5M acetic acid
Ethanol	100% and 70% (diluted with distilled water)

2.1.10) Reagents and solutions for large scale plasmid preparation using Midiprep QIAGEN Kit (Crawley, West Sussex, UK).

Buffer P1	50mM Tris HCl (pH 8.0), 10mM EDTA, 100ng/ml Rnase A
Buffer P2	0.2 M NaOH, 1%(w/v) SDS
Buffer P3	3M potassium acetate (pH 5.5)
Buffer QBT	750mM NaCl, 50mM MOPS (pH 7.0), 15% isopropanol, 0.15% triton X-100
Buffer QC	1M NaCl, 50mM MOPS (3-(N-Morpholino) propanesulfonic acid (pH 7.0), 15% isopropanol
Buffer QF	1.25 M NaCl, 50mM Tris HCl (pH 8.5), 15% isopropanol

2.1.11) Reagents and solutions for transfection

2.1.11A) LIPOFECTAMINE Plus Transfection Reagent (Invitrogen, Paisley)

Lipofectamine Plus Transfection Reagent package consists of LIPOFECTAMINE ™ Reagent and the proprietary Plus Reagent prepared for pre-complexing DNA.

2.1.11B) FuGENE™ 6 Transfection Reagent (Roche, Lewes)

FuGENE Reagent is a proprietary blend of lipids (non-liposomal formulation) and other compounds in 80% ethanol, filter sterilised.

2.1.12) Antibodies for immunofluorescence

Rabbit anti HBcAg polyclonal IgG	Zymed, San Francisco, CA
FITC conjugated goat anti rabbit IgG	Sigma and used as secondary detection Ab at a working dilution of 1:64

2.1.13) Commom Reagents

Manufacturer	Chemicals
Beecham	Ampicillin
Sigma	Aphidicolin
Melford	X-gal
Gibco BRL	IPTG (isopropyl-β-D) thiogalactopyranoside

2.2 METHODS

2.2.1 HBV DNA Extraction, PCR, Sequencing and Phylogenetic analysis

2.2.1A) High Pure Viral Nucleic Acid Kit (Roche Diagnostics, Lewes, UK).

For the Roche kit, 200 ul of working solution (binding buffer supplemented with poly (A) carrier RNA) and subsequently 40 ul of 20 mg/ml proteinase K were added to 200 ul of serum in a sterile Eppendorf tube, mixed and incubated for 10 min at 70° C. After the incubation, 100 ul of isopropanol was added. The filters and collection tubes were combined and the samples pipetted into the upper reservoir followed by centrifugation for 1 min at 8,000 x g and flowthrough discarded. The filter was washed twice with the wash buffer and the flowthrough discarded after each wash. Finally, centrifugation for 10 seconds at full speed removed the entire residual wash buffer. Collection tubes were discarded and clean nuclease-free 1.5 ml tubes were used to collect the eluted DNA in 50 ul of elution buffer. The eluted DNA was stored at -20°.

2.2.1B) PCR amplification of the extracted DNA

Hot start PCR was performed using a hemi-nested protocol and antibody to Taq polymerase to amplify the core gene of HBV. Five ul of extracted DNA was amplified in 50 ul solution containing: 1.25 U Taq polymearse (Gibco, Paisley, UK), 2.5 ul TaqStart antibody (Clontech Laboraties Inc., Palo Alto, Ca, USA), 0.25 mM dNTPs (Pharmacia, St. Albans, UK), 10X PCR buffer (supplied with Taq polymerase), and 25 pmol of each primer (C1: sense and C4: antisense) (Table 2.1A); cycled for 5 cycles of 95° C for 1 min, 55 ° C for 1 min, and 72° C for 90 seconds followed by 35 cycles. One microliter of first round PCR product was then re-amplified in the same solution as above except for heminested primers, (C3a: sense and C4: antisense, table 1). For 5 cycles of 95° C for 1

min, 55 ° C for 75 seconds, and 72° C for 90 seconds followed by 25 cycles with the denaturation temperature reduced to 90° C. A nested protocol was used for the amplification of the surface region. Primers S1 and S2 followed by S6c and S7D were used to amplify the surface gene (Table 2.1A). The reaction mixture and program cycles of Biometra TRIO Thermoblock, were similar to those used in C gene amplification. Amplification of S gene was employed to confirm the PCR or sequencing results of C gene.

2.2.1C) Agarose gel electrophoresis

Agarose gel electrophoresis was used to confirm the right size of our amplified PCR products by visualising the gel on an ultraviolet transilluminator. Also, it was used to check linearised plasmids (after enzyme digestion and before ligation) and restriction digests of miniprep DNA. Gels were prepared by adding 1 g agarose to 100ml 1X TBE buffer. The solution was boiled until dissolved and left to cool. Then, 50ul ethidium bromide (1mg/ ml) was added before pouring the gel. One ul of agarose gel loading buffer was used to each sample before loading, followed by running the gel at 80 V in 1X TBE buffer for 25 min.

2.2.1D) Purification of the PCR products by High Pure PCR Product purification kit (Roche Diagnostics, Lewes, UK)

250ul of binding buffer was added to 50ul of PCR reaction product and mixed well. The mixture was poured onto the High Pure filter tube and centrifuged at 13,000rpm for 30 seconds. The filter tube was washed twice. Finally, the filter was inserted in RNase free 1.5 ml tube and dH_20 was applied to elute the DNA (higher volumes of dH_20 for elution are preferred as it increases the elution efficiency).

2.2.1E) DNA sequencing

2.2.1E I) DNA purification and sequencing

DNA was firstly purified by using the previously mentioned kit. Sequencing of the C gene was performed either directly from the purified PCR products or after C gene cloning (by mutagenesis, see section 2.2A) using an automated sequencer (ABI Prism, 377 DNA sequencer, Applied Biosystem, Perkin Elmer) according to the manufacturer's instruction. The HBsAg genotype of the sequences was defined by substitutions between codons 122 and including 160. The reaction mixture was prepared by adding 2-4 pmol of 1pmol/1ul of each primers (core: C3a and C4; surface: S6C and S7D) (Table 1.2A) to 1-2 ul of PCR products or 400-800 ng of cloned DNA up to 6 ul of reaction mixture by adding dH_20. Because of unavailability of DNA materials, we did not sequence the S gene of Indian samples.

2.2.1E II) Sequence data analysis

The sequence Navigator software program (Applied Biosystem, Cheshire, UK) was used to analyse our sequence data. The sequences were aligned and the consensus sequence was determined for sequences that have the same subtype/ genotype.

After allocating a sequence to an HBV genotype by analysis of the S gene, the core gene amino acid/nucleotide variations that were found uniquely in that subtype or genotype were recorded. Variants found in a majority of HBcAg sequences (> 50%) from a particular geographical area were confirmed with HBcAg sequences from the same area obtained from EMBL, GenBank and NCBI (Table 3.1.2). Sequences were only selected from databases if the subtype, country of origin and serological picture were provided. Finally, geography-specific nucleotide and amino acid variations in each genotype were

identified. To qualify, they had to be both unique to a geographical region and found in a majority of sequences from that region. Sequences studies from chapter 3.1 have been submitted to Genbank, numbered from AF323463 to AF323470 and from AF324066 to AF324148. Those sequences obtained from the Pacific (chapter 3.2) study are numbered from AY269035 to AY269047, from AY269050 to AY269061 and from AY269064 to AY269086.

2.2.1F) Phylogenetic Analysis

Sequences of core and surface genes were aligned using the BioEdit Package version 5.0.9, and a neighbor joining phylogenetic tree constructed using the Treecon Package employing a Kimura distance matrix. Associations were tested by bootstrap re-sampling analysis using 100 replicates. Associations with a bootstrap value of greater than 70% were deemed significant.

2.2.2 Molecular cloning and transfection methods

2.2.2 A) ExSite™ PCR-Based Site-directed mutagenesis

Site-directed mutagenesis was performed by a PCR method using oligonucleotides (Table 2.1B) to revert mutated sequences back to the original sequence, according to manufacturer's instructions (Figure 2.1) (Stratagene, La Jolla, CA). PCR reactions were set up for two separate reactions: control and sample reactions. Oligonucleotides (primers 1 and 2) and plasmid template control (an ampicillin-resistant pWhitescript™ 4.5 kb control reaction) were included in this kit.

2.2.2A I) Reaction Setting Up

The control reaction contained the following:

PWhitescript	1 ul (0.5 pmol, ~1.5 ug)

10x mutagenesis buffer	2.5 ul
dNTPs mix	1ul of 25 mM
Primer#1 (phosphorylated)	2 ul (15 pmol, ~150 ng)
Primer#2	2 ul (15 pmol, ~150 ng)
ddH_2o	15.5 ul (to a final volume of 24 ul)

The sample reaction contained the followings:

DNA	0.5 pmol of template
10x mutagenesis buffer	2.5 ul
dNTPs mix	1ul of 25 mM
Primers (side-directed mutagenesis)	15 pmol of each primer
ddH_2o	to a final volume of 24 ul

Then 1 ul of ExSite DNA polymerase blend (5 U/ul) added to each reaction tube, overlaid with 20 ul of mineral oil.

2.2.2A II) Cycling Parameters

As shown in table 2.2, the following cycling parameters were used for PCR reaction in the site-directed mutagenesis procedure. Because of PCR negativity of some samples, different annealing temperatures between 50° to 56° were applied.

2.2.2A III) Digestion and polishing the product

Following the PCR, the reactions were cooled on ice for 2 min. Nonmutated plasmids were subsequently digested by Dpn I. The addition of 1 ul (10U/ul) of Dpn I, 0.5 ul (2.5 U/ul) of P*f*u DNA polymerase (to remove extended

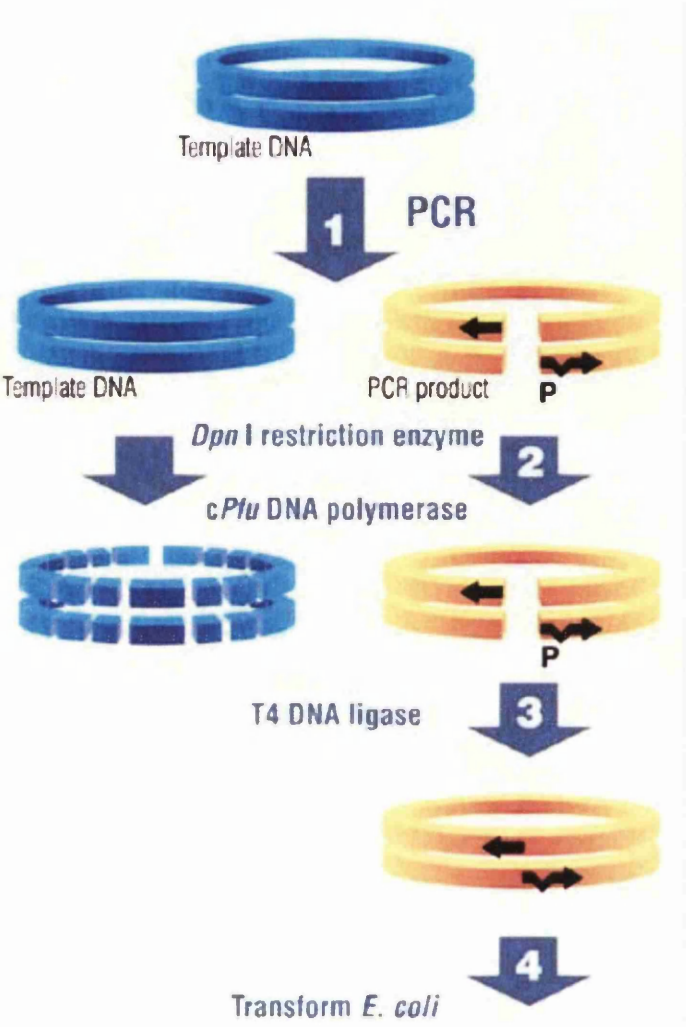


Figure 2.1 A diagram on the basis of ExSite™ PCR-Based Site-directed mutagenesis. (Taken from: http://www.stratagene.com/displayproduct.asp).

The ExSite method uses increased template concentration. The resulting mixture of template DNA, newly synthesized DNA and hybrid parental/newly synthesized DNA is treated with *Dpn* I and *Pfu* DNA polymerase. *Dpn* I digests the in vivo methylated parental template and hybrid DNA, and *Pfu* DNA polymerase polishes the ends to create a blunt-ended PCR product. The end-polished PCR product is then intramolecularly ligated together and transformed into *E. coli.*

Table 2.2: PCR cycling parameters

Step	Cycle	Temperature	Time
1	1	94°	4min
		50°	2min
		72°	2min
2	8	94°	1min
		50°	2min
		72°	1min
3	1	72°	5min

3' extension at the end of PCR products) below the mineral oil. Digestion was carried out by incubation at 37° for 30 min. The reaction was then incubated for 30 min at 72°. The Dpn I- cloned P*fu* DNA polymerase-treated DNA was then ligated at 37° for 1 h with 1 ul (4U/ul) of T4 DNA ligase.

2.2.2A IV) Transformation

The Epicurian Coli XL1-Blue supercompetent cells (kept at –70°) were thawed gently on ice for a few min. 80 ul of cells aliquoted to a Falcon® 2059 polypropylene tube, 2 ul of ligase-treated DNA added to cells, swirled, and incubated on ice for 30 min. For control transformation, 0.1 ng of pUC18 transformation control plasmid added to cells, swirled and incubated for 30 min on ice. Then, the tubes were heat pulsed for 45 s at 42° and then placed on ice for 2 min. Then the entire volume of transformed cells plated on LB-ampicillin agar plates that have been spread with 20 ul of 10% (w/v) X-gal prepared in dimethylformamide (DMF) and 20 ul of 100 mM IPTG prepared in filter-sterilized dH_2O, incubated overnight (> 16 h).

2.2.2B) Small Scale Plasmid Preparation (Mini-prep)

After checking the control plasmid (which should appear as blue on agar plates containing IPTG and X-gal), white colonies of transformed bacteria were inoculated into 2-3 ml of LB broth containing ampicillin at 100ug/ml and incubated in a shaker at 37° C overnight. One ml of the overnight culture was transferred to a 1.5 ml eppendorf tube and centrifuged at 13,000 rpm for 30 seconds. The pellet, after decanting the supernatant, was resuspended in 100ul of ice cold solution (I) and left at RT for 5 min. 200ul of freshly prepared solution (II) was added, mixed gently and left on ice for 2-3 min. 150ul of solution (III) was added, mixed by few inversions, left on ice for 2-3 min. 150ul of phenol/chloroform was added to the mixture, vortexed and centrifuged at 13,000rpm for 3-5 min. The aqueous phase was transferred to a fresh tube, 800ul of 100% ice cold ethanol was added to precipitate the plasmid DNA and centrifuged for 10min. The pellet was then washed with 70% chilled ethanol, air dried and finally resuspended in 50ul dH_20 containing RNase at 20ug/ml to remove any contaminating RNA.

2.2.2C) Restriction enzyme digestion

Restriction digestion of plasmid DNA was carried out to confirm successful cloning. The standard mixture consisted of 1.5ul 10X restriction buffer (B), 0.2ul (2 U) of each restriction enzyme (EcoR1 and Hind-III), 2.5-5ul plasmid DNA and dH_20 to a final volume of 15ul. The digest mixture was then incubated at 37° C for 2-3 hr.

2.2.2D) Large Scale Plasmid DNA Preparation

Large scale plasmid preparations were performed to purify plasmid DNA, using the QIAGEN midi kit. 0.5 ml of the overnight culture, of confirmed clones, was added to 50ml L-broth containing 100ug/ml ampicillin and incubated overnight at 37° C. The

culture volume was divided into two equal parts which were spun at 3,000rpm for 15 min at 4° C. One of the cell pellets was stored at –20 ° C, while the second was resuspended by vigorous vortexing in 4ml of buffer P1. 4ml of buffer P2 was added and mixed gently by several inversions. After 5 min incubation at RT, 4ml of chilled buffer P3 was added and mixed gently. The lysate was poured into the QIAfilter cartridge with screw cap on and incubated at RT for 10 min. At the meantime, QIAGEN-tip 100 was equilibrated with 4ml of buffer QBT. The lysate filtrate was then transferred from QIAfilter to the equilibrated QIAGEN-tip and left to drip by gravity flow. After all the lysate filtrate had dripped through the QIAGEN-tip, the resin tip was washed twice with wash buffer QC. Elution of the plasmid DNA from the resin tip was finally achieved by addition of 5ml elution buffer QF. The plasmid DNA was precipitated by the addition of 3.5ml isopropanol followed by centrifugation at 11,000 rpm (Sorvall SM24 rotor) for 30 min at 4° C. The pellet was then washed with 70% ethanol, air dried and resuspended in 100ul dH_20. Finally, the plasmid DNA concentration was determined by measuring the absorbance at 260 nm.

2.2.2E) Cell culture

It is known that the distribution of expressed HBcAg antigen is cell cycle dependent. Sequential thymidine and aphidicolin blocking steps produced monolayers of synchronised Cos-7 cells. Cos-7 cells were seeded at a density of $3X10^4$ cells per well in a 24-well plate in growth media 12 hours before treatment with cell cycle blocking agents (Fig 2.2). Then, media containing 2mM thymidine was substituted, and 12 hours later cells were washed twice by PBS and refed with normal media. Transfections were carried out after cell release, using FuGENE-6 and/or Lipofectamine Plus Transfection Reagent

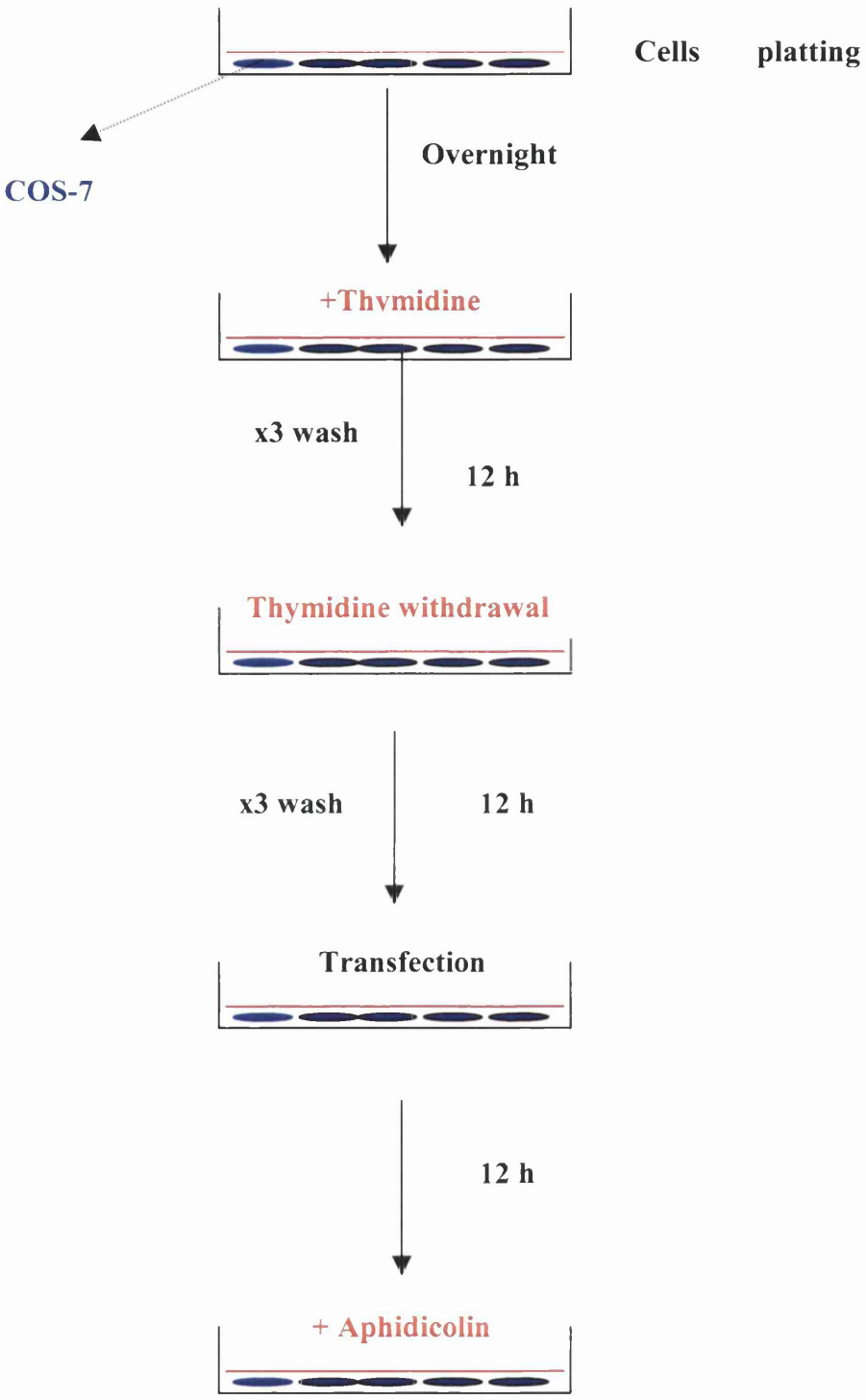


Fig 2.2. A diagram showing sequential thymidine and aphidicolin blocking stepsfor cell cycle synchronisation.

methods according to the manufacturer's instructions (see bellow). A further 13 hours later, cells were washed by PBS twice and media containing 2.5 μg/ml aphidicolin was added, then incubated for further 36 hours. Aphidicolin blocks cells in the G1/S phase of the cell cycle, in which HBcAg expression is predominantly nuclear (Yeh et al., 1993).

2.2.2F) Transfection

2.2.2F I) FuGENE ™ 6 Transfection Reagent

In this method, The DNA was transfected into subconfluent monolyers of COS-7 cells on 13 mm coverslips. 1.5 ul of FuGENE was added to 25 ul of Optimum 1 reduced-serum medium (Life Technologies, Paisley, UK), followed by 300-500 ng of DNA, incubated for at least 15 min at RT. Then, this mixture was added to the adherent subconfluent COS-7 cells in 24 wells plate each well contained 1 ml DMEM, incubated at 37° C at various timescales according to cell cycle blocking protocol.

2.2.2F II) LIPOFECTAMINE™ PLUS Reagent

25 ul of Optimum reduced-serum medium (Life Technologies, Paisley, UK), inserted into a clean sterile 1.5 ml eppendorf tube followed by 200-300 ng of plasmid DNA, mixed and then 4 ul of PLUS Reagent added, mixed again and incubated at RT for 15 min. In a separate 1.5 tube 1ul of LIPOFECTAMINE™ Reagent added to 25ul of Optimum, mixed. The contents of pre-complex DNA combined to the diluted LIPOFECTAMINE™ Reagent tube, mixed and incubated at RT for 15 min. This mixture added to each well containing fresh medium (0.2 ml) incubated at 37° for 3 h.

After 3 h, the medium volume increased to normal volume and incubated at 37° in different timescale according to cell cycle protocol. Different reagents and volumes used in this protocol summarized below.

DNA (ug)	0.2-0.3
PLUS Reagent (ul)	4
Optimum (ul)	25
LIPOFECTAMINE Reagent (ul)	1
Transfection Medium (ml)	0.2
Transfection Volume (ml)	0.256

2.2.2G) FACS Analysis

Cell cycle status of cultured cells was determined by flowcytometry of permeabilised cells that were stained with propidium iodide (PI). Cos-7 cells were dissociated with EDTA at the appropriate time post-transfection and resuspended in PBS. Samples were centrifuged at 1500 g for two minutes, then cells were fixed with 70% ethanol. After fixation on ice for 20 minutes, cells were centrifuged, washed once with PBS and centrifuged again. Pelleted cells were resuspended in 1 ml of a solution contained 100X Triton, PI, and RNase. After 30 minutes incubation at room temperature, the total DNA contents was analysed by Becton Dickinson FACScan instrument with CellQuest software (Becton Dickinson, San Jose, CA).

2.2.2H) Immunofluorescence

Transfected cell monolayers on coverslips were washed twice with PBS. The cells were then fixed with ice cold methanol (at RT for 10 min), washed three times with PBS and

permeabilised with 0.5% triton-X100 in PBS. The permeabilised cells were rinsed twice with 0.05% Tween in PBS and incubated with rabbit anti-core polyclonal IgG antibody at RT in a moist dark box. After washing, the cells were further incubated with FITC conjugated goat anti-rabbit IgG as secondary antibody at 1:64 dilution, incubated for 45 min. After three washes, the coverslips were dried carefully and mounted on glass slides with a drop of Citifluor, a glycerol/PBS solution.

2.2.2F) Confocal Microscopy

Localisation of core protein was studied using confocal microscopy. Then cell samples were examined in a Zeiss LSM510 confocal microscope with laser excitation line at 488nm. The microscope was a Zeiss Axiplan with either a 40X or 63X oil immersion objective lens. Data sets were processed with the Laser Scanning Microscope (LSM 510) software.

CHAPTER 3 - Core Gene Variability, Geography and Ethnicity

3.1 Aims of study

Although "core aa substitution" was reported in many studies, some of these variants are actually the consensus sequence from other genotypes. However, in view of the heterogeneity of HBV strains, it is difficult to determine whether core gene substitutions, recorded in the literature (in addition to those characteristic of genotype), only once represent naturally occurring variants or an immune-selected mutation in individuals. The aims of this study were: to characterise the core gene variability in diverse geographic regions and ethnic groups; to determine subtype/genotype-specific variants; and to specify country/ethnic-specific variant. The current work can be divided into two parts. In the first, Sera from 91 HBsAg positive patients from 9 countries, including 49 samples from California (39 from Asian origin persons living in California and 10 from Caucasians), 12 from Italy, and 10 each from Scotland, India, and the Pacific (Table3.1.1) were analyzed by sequencing the core gene. All were positive for HBsAg with a variety of clinical features and HBeAg/anti-HBe serological status. The Asian-origin samples from California were previously sequenced for a pre-core variability study (results not shown). These findings then were compared to sequences of known geographic origin in the database to confirm the associations.

Unique sequences were obtained from the Pacific group at both aa and nucleotide level, (first part of this chapter), leding us to more investigation on core gene variability in this area. We selected randomly 58 samples from HBV-infected patients from 4 different islands. HBV isolates from C and D genotypes were identified. Genotype C is predominant in Vanuatu, Fiji and Tonga, while D is the dominant genotype in Kiribati. HBs Ag positive samples were collected from four Pacific islands (Table 3.2.1): Kiribati (14), Vanuatu (14), Fiji (15) and Tonga (15). Geographically, these islands are located as follows: Kiribati in Micronesia; Fiji and Vanuatu in Melanesia; and Tonga in Polynesia (Fig 3.2.1). They were chosen from a

previous study of surface variability and correlation with the history of human populations and immigration patterns of the Pacific region (Basuni 2001).

PCR was employed to amplify the C and S genes, followed by direct sequencing using a variety of forward and backward primers to eliminate sequence variation caused by possible errors of Taq polymerase. The nucleotide sequences were then edited and assembled using Sequence Navigator software. We already knew the genotype as this had previously been characterised by surface gene sequencing. Comparative sequence analysis to detect important nucleotide changes and motifs was performed using the BioEdit Package version 5.0.9. Tables showing alignments of amino acid and nucleotide sequences are shown in the appendices at the back of this thesis.

3.2 Results

3.2-a) Phylogenetic Analysis

Figures 3.1.1, A and B show that sequences grouped into 6 major clusters. Four of these were occupied by genotypes A to D. Pacific sequences composed 2 additional clusters (in both C and S trees), one in genotype C and one in D. Compared to the rest, the most homogeneity was observed in Pacific sequences. In genotype D strains, most of the Indian sequences comprised a sub-cluster in the C tree. Similarly, a majority of the strains found in Scotland and the USA (genotype A), were homogenous in both C and S trees. The phylogenetic trees constructed based on S and C genes revealed that the topological features of all strains except of A4, A5, I54 and I113 in the phylogenetic trees were identical. The C gene of A5 belonged to genotype B, whereas, the S gene belonged to genotype C (the Pacific adrq-). Similarly, strains I54 and I113, based on core sequencing had genotype D, whereas, the S genes were genotype A. In one genotype D cluster, there were 2 Western-derived samples amongst Indian sequences. These samples showed homology at the amino acid level; however, there was a substantial amount of synonymous nucleotide variation (Tables A.I.1 and A.I.2).

3.2-b) Genotype/Subtype-specific Nucleotide Substitution

At the nucleotide level, it was possible to identify unique variants for subtypes adw2 (belonging to genotypes A and B), ayw1, ayw2, adr and adrq- (there was no ayr subtype in the sequences studied). Overall, 37 genotype specific nucleotide substitutions in 26 positions were found (Tables 3.1.3 and A.I.2), of which 16 were silent (synonymous) and 21 were missense (non-synonymous). Of 37 variations, it was possible to define 9 unique nucleotide variants for specific subtypes: C_{221} and G_{342} for ayw2; C_{342} for ayw3 strains; T_{66}, G_{75} and G_{175} for adrq-; and C_{289} for adrq+. A_{454} was characteristic for adw2 and A_{404} was characteristic for ayw1, both belonging to genotype B. Our results concurred with those of sequences in databases.

3.2-c) Amino acid Substitutions and Genotype/Subtype

The amino acid residue at 9 codons (27, 59, 67, 74, 83, 87, 91, 97 and 116), and an insertion at amino acid position 152, allowed allocation of a specific subtype (Table 3.1.4). Of these 9 residues, the following were characteristic for genotypes and their corresponding subtypes: N74 and N87 in adw2 genotype A (adw2); F97 and I116 in genotype D (ayw2 and ayw3 strains); V59 and I91 in genotype C strains (for adrq-). All sequences which encoded adw2 (genotype A) had a 2-codon insertion corresponding to amino acid 153 and 154 in the core gene, independent of origin. The most divergent residues were amino acids 74 and 91 with 5 and 3 possible variant amino acids respectively. Tables A.I.1 and A.I.2 show amino acid and nucleotide alignment of all HBcAg sequences in the countries studied, respectively. Most amino acid substitutions involved the HLA-restricted CTL, T helper and B cell epitopes, irrespective of HBV subtypes (Table 3.1.4; see chapter one table 1.3).Genotypes A (adw2 subtype) and D sequences had 5 amino acid substitutions in common at residues 12, 27, 67, 83 and 91, which were not present in other strains. Distinguishing adw2 (genotype A) and ayw2 was possible by examination of residues 74, 87, 97, and 116 (Table 3.1.4), confirming

that substitutions in these positions alone explain d/y variation in the C gene. The most heterogeneity was seen in ayw2 strains derived from Indian, Italian, Pacific island and Scottish patients (Table A.I.1). Residue 74 showed the most discrepancy between countries in which ayw2 is common: T in Pacific, V in Scottish and Italian, and G in Indian. At the amino acid level, the core gene sequences belonging to adw2, ayw1 (both belonging to genotype B) and adrq+ did not show any unique amino acid substitution (however, see nucleotide results). Interestingly, with the exception of two samples, all strains with 97L (instead of I) were HBeAg negative or had a pre core stop codon, A_{1896} (results not shown).

3.2-d) Country-Specific Changes

A specific motif, consisting of S12, I27, N67, S74, E83 and V91, was observed in a majority of samples from South-east Asia and the Pacific irrespective of HBV genotype, but not in core sequences from other regions studied (Table 3.1.5A). Similarly, Western-derived sequences (USA-Caucasian, Scotland and Italy) shared a specific motif at the same residues, but with different amino acids: T12, V27, T67, N/V74, D83 and T91 (Table 3.1.5A). This pattern seemed slightly different at amino acid position 74 in which all variation depended on genotype: N74 in genotype A and V74 in genotype D. This finding was strengthened by analysis of the nucleotide sequence (Table A.I.1). In both populations, 19 unique nucleotide changes were found, of which 8 were missense changes for 6 amino acid residues (Table 3.1.5-B). There were nine Indian and nine Italian strains of genotype D, subtype ayw2. One sequence (represented by six examples) dominated in each country; the remaining three sequences were represented by one example each (Table A.I.1). Further, the five Scottish strains of genotype D, subtype ayw2 or ayw3, were identical to the majority strains in India or Italy, while all the Pacific-derived strains of genotype D were unique and identical.

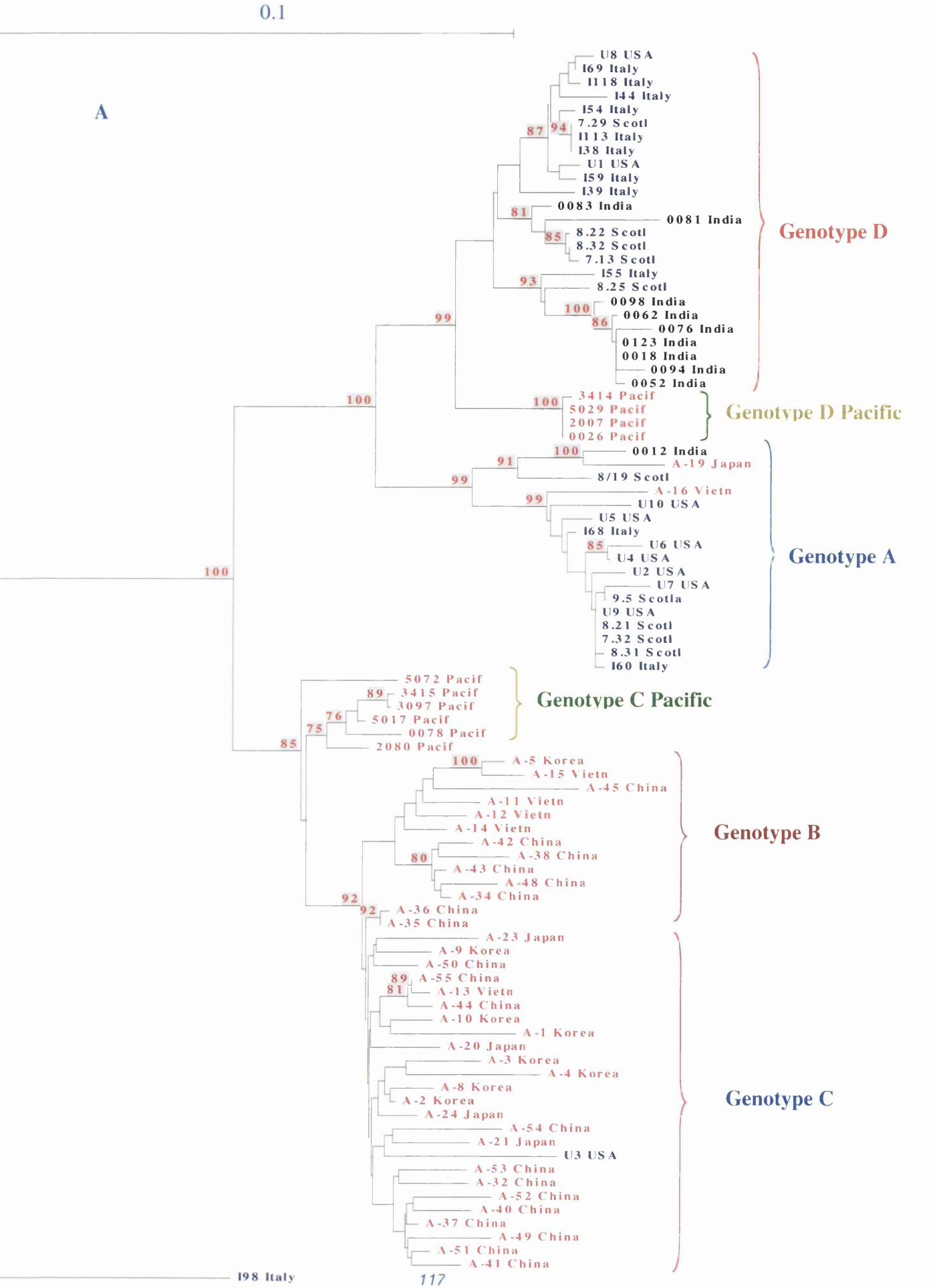
Seven amino acid positions allowed definition of a particular country's genotype D sequence. In the Pacific all sequences belonging to ayw2 contained A35, D40, and T74; V59 and I91

was observed in a majority of adrq-strains from that region. I80 was seen in Indian sequences and G87 in Korea (Table 3.1.6). The most heterogeneity was seen in ayw2 strains derived from Indian, Italian, Pacific island and Scottish patients. Residue 74 showed the most discrepancy between countries in which ayw2 is common: T in Pacific, V in Scottish and Italian, and G in Indian (Table A.I.1).

However, there were additional synonymous nucleotide substitutions that were characteristic for specific countries in our study (Table 3.1.6). For example, in a majority of samples from India, we observed T_{12}, C_{55}, T_{57}, A_{111}, G_{132}, A_{238}, G_{255} and T_{291}. In all strains belonging to ayw2 from the Pacific region, we observed T_{66}, G_{75}, G_{103}, G_{105}, T_{120}, A_{189}, C_{221} and T_{255}. A_{342} and C_{289} were characteristic for Italian and Korean sequences respectively.

Finally, subtype-related specific nucleotide variations were observed in a majority of samples from specific countries. In keeping with the amino acid sequence, the ayw2 subtype contained the most nucleotide divergence (Table 3.1.7): 16 nucleotide substitutions were unique in ayw2 sequences from India, 10 from Pacific, and 2 from Italy. The same heterogeneity was observed in adrq+ and adrq- strains (Table 3.1.8): all Pacific strains with adrq- subtype showed 5 specific changes. A_{258} and C_{289} were unique in adr strains from Chinese and Korean sequences respectively.

Fig. 3.1.1- Neighbour joining phylogenetic trees of core (A) and surface (B) genes sequences from 91 HBcAg samples. Note: C and S gene trees rooted with sample I98 and U8, respectively. The figure shows bootstrap values of ≥70% and scale denotes percent diversity. Coding numbers indicate samples that have been analysed in the figure. Because of unavailability of DNA materials, we did not sequence the S gene of Indian samples. Scotl=Scotland; Pacif=Pacific and Vietn=Vietnam.


0.1
A
U8 USA
I69 Italy
I118 Italy
I44 Italy
I54 Italy
7.29 Scotl
I113 Italy
I38 Italy
U1 USA
I59 Italy
I39 Italy
0083 India
0081 India
8.22 Scotl
8.32 Scotl
7.13 Scotl
I55 Italy
8.25 Scotl
0098 India
0062 India
0076 India
0123 India
0018 India
0094 India
0052 India
Genotype D
3414 Pacif
5029 Pacif
2007 Pacif
0026 Pacif
Genotype D Pacific
0012 India
A-19 Japan
8/19 Scotl
A-16 Vietn
U10 USA
U5 USA
I68 Italy
U6 USA
U4 USA
U2 USA
U7 USA
9.5 Scotla
U9 USA
8.21 Scotl
7.32 Scotl
8.31 Scotl
I60 Italy
Genotype A
5072 Pacif
3415 Pacif
3097 Pacif
5017 Pacif
0078 Pacif
2080 Pacif
Genotype C Pacific
A-5 Korea
A-15 Vietn
A-45 China
A-11 Vietn
A-12 Vietn
A-14 Vietn
A-42 China
A-38 China
A-43 China
A-48 China
A-34 China
A-36 China
A-35 China
Genotype B
A-23 Japan
A-9 Korea
A-50 China
A-55 China
A-13 Vietn
A-44 China
A-10 Korea
A-1 Korea
A-20 Japan
A-3 Korea
A-4 Korea
A-8 Korea
A-2 Korea
A-24 Japan
A-54 China
A-21 Japan
U3 USA
A-53 China
A-32 China
A-52 China
A-40 China
A-37 China
A-49 China
A-51 China
A-41 China
Genotype C
I98 Italy

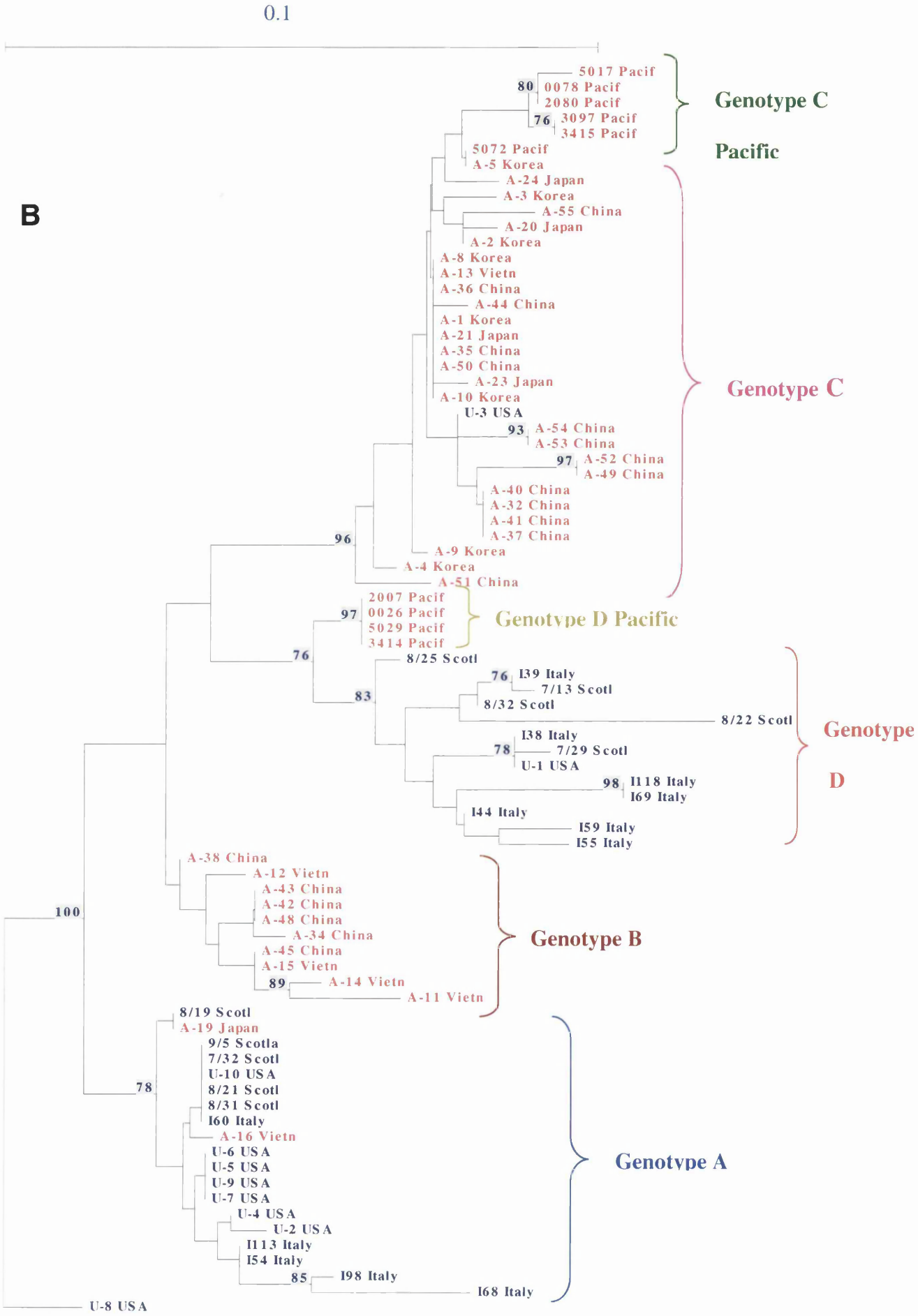

0.1
B
5017 Pacif
80
0078 Pacif
2080 Pacif
76
3097 Pacif
3415 Pacif
5072 Pacif
A-5 Korea
Genotype C
Pacific
A-24 Japan
A-3 Korea
A-55 China
A-20 Japan
A-2 Korea
A-8 Korea
A-13 Vietn
A-36 China
A-44 China
A-1 Korea
A-21 Japan
A-35 China
A-50 China
A-23 Japan
A-10 Korea
U-3 USA
93
A-54 China
A-53 China
97
A-52 China
A-49 China
A-40 China
A-32 China
A-41 China
A-37 China
96
A-9 Korea
A-4 Korea
A-51 China
Genotype C
2007 Pacif
97
0026 Pacif
5029 Pacif
3414 Pacif
Genotype D Pacific
76
8/25 Scotl
76
I39 Italy
7/13 Scotl
83
8/32 Scotl
8/22 Scotl
I38 Italy
78
7/29 Scotl
U-1 USA
98
I118 Italy
I69 Italy
I44 Italy
I59 Italy
I55 Italy
Genotype
D
A-38 China
A-12 Vietn
A-43 China
A-42 China
A-48 China
A-34 China
A-45 China
A-15 Vietn
89
A-14 Vietn
A-11 Vietn
100
Genotype B
8/19 Scotl
A-19 Japan
9/5 Scotla
7/32 Scotl
U-10 USA
78
8/21 Scotl
8/31 Scotl
I60 Italy
A-16 Vietn
U-6 USA
U-5 USA
U-9 USA
U-7 USA
U-4 USA
U-2 USA
I113 Italy
I54 Italy
85
I98 Italy
I68 Italy
Genotype A
U-8 USA

Table 3.1.1: Origin of 91 HBsAg positive sera that were used as the source for HBcAg sequencing.

Country		Number of Samples
South-East Asia		
Californian samples	China	20
	Korea	8
	Japan	5
	Vietnam	6
	Pacific	10
Caucasian		
	USA	10
	Scotland	10
	Italy	12
	India	10
Total		91

Table 3.1.2: Identification of core sequences obtained from databases for comparison.

Accession Number	Identification	Origin	Subtype	Genotype
M57663	pFDW294	Philippines	adw2	A
X02763	PHBV3200	USA	adw2	A
X51970	HBV991	Germany	adw2	A
X70185	HBV-A938	Germany	adw2	A
AF297624	Isolate 656	Africa	adw2	A
V00866	PHBV933	USA	adw2	A
Z35717	pHB614	Poland	adw2	A
X97850	Patient-4	China	adw2	B
X97851	Pateint-6	China	adw2	B
AB033554	RTB299	Indonesia	adw2	B
AB033555	PAD744	Indonesia	adw2	B
AF121249	AF121249	Vietnam	adw2	B
AF282917	HBV-B_1	China	adw2	B
AF282918	HBV-B_2	China	adw2	B
AB031261	HBV Vie A-2	Vietnam	adw2	B
AB031263	HBV Vie A-5	Vietnam	adw2	B
AB031264	HBV Vie F-1	Vietnam	adw2	B
AB031262	HBV Vie A-3	Vietnam	adr	C
AB031265	HBV Vie F-2	Vietnam	adr	C
AB033553	SK619	Indonesia	adw	C
D50489	HPBA11A	Japan	adr	C
L08805	HPBETNC	Japan	adr	C
D16666	HPBE88A	Japan	adr	C
S75184	S75184	Japan	adr	C
AF241411	8290	Vietnam	adr	C
AB031260	HBV VieA-1	Vietnam	adr	C
X52939	HBV prex	China	adr	C
M38454	pADR-1	China	adr	C
M38594	M38594	Korea	adr	C
M38636	pHBV107	Korea	adr	C

X04615	pYRB259	Japan	adr	C
X14193	pADRm	Korea	adr	C
D12980	SRADR	Japan	adr	C
D00630	pHBV330	Japan	adr	C
X01587	pHBVadr4	Japan	adr	C
D00331	pAK66	Japan	adr	C
X75656	HHVCCHA	Polynesia	adrq-	C
X75665	HHVBC	New Caledonia	adrq-	C
AB033557	pIWK146	Indonesia	adw	C
AF121240	HBV/94-11066	Vietnam	ayw	D
AF121242	HBV/98-1218	Vietnam	ayw	D
AB033558	JYW796	Japan	ayw	D
AB033559	JYW310	Papua New Guinea	ayw	D
X65258	HBVAYWCI	Italy	ayw	D
X59795	HBVAYWMCG	Italy	ayw	D
X02496	pHBV320	Latvia	ayw	D
AJ131956	AJ131956	Germany	ayw	D
L27106	HPBMUT	Israel	ayw	D
X72702	HBVORFS	Germany	ayw	D
X85254	HBVPRESS12	Italy	ayw	D

Note: Each sequence is listed under accession number.

Table 3.1.3: Unique C-gene nucleotide differences which permit subtype classification.

Nucleotide	Subtype (genotype)						
Position	adw2(A)	adw2(B)	ayw1(B)	adr(C)	adrq-(C)	ayw2(D)	ayw3(D)
66	C	C	C	C	T	C	C
75	T	T	T	T	G	T	T
81	C	T	T	T	T	A	A
159	T	G	G	A	A	T	T
174	C	T	T	T	T	A	A
175	A	A	A	A	G	A	A
180	C	G	G	G	G	T	T
204	A	A	A	G	G	A	A
221	A	G	G	G	G	T/G/C	T
249	T	A	A	A	A	C	C
260	A	G	G	G	G	G	G
289	A	A	A	A /C	A	T	T
324	T	T	T	T	T	C	C
342	T	T	T	T	T	G/A	C/A
346	C	C	C	C	C	A	A
390	A	T	T	C	C	A	A
393	C	A	A	T	T	T	T
396	T	T	T	C	C	T	T
404	C	C	A	C	C	C	C
454	C	A	C	C	C	C	C
454-459 insertion	CGG GAC	-	-	-	-	-	-

Note: Underlined numbers indicate the locations of the missense nucleotides. Nucleotides are represented by single letters and numbered from the beginning of HBcAg.

Table 3.1.4: The most frequent differences in HBc Ag amino acid sequences between subtypes.

Subtype/ Genotype	CD8	CD4				CD8			
	27	**59**	**67**	**74**	**83**	**87**	**91**	**97**	**116**
adw2/A	V	I	T	N	D	N	T	I	L
adw2/B	I	I	N	S	E	S	V	I	L
ayw1/B	I	I	N	S	E	S	V	I	L
adr/C	I	I	N	S	E	S	V	I	L
adrq-/C	I	V	N	S	E	S	V/I	I	L
ayw2/D	V	I	T	G/T/V	D	S	T	F	L/I
ayw3/D	V	I	T	V	D	S	T	F	I

Note: CD4 = CD4 recognized epitope. CD8 = CD8 recognized epitope. Each variation was found in a majority of strains from each subtype and the combination was unique for that subtype. Numbering of amino acid in bold is from the beginning of HBcAg.

Table 3.1.5(A)- Specific amino acid motifs in different ethnic groups.

Country	China		Korea	Japan	Vietnam	Pacific	USA	Italy		Scotland	
Sub/Genotype	adr/C	adw2/B	adr/C	adr/C	ayw1/B	adrq-/C	adw2/A	adw2/A	ayw2/D	adw2/A	ayw2/D
Amino Acid Position											
12	S	S	S	S	S	S	T	T	T	T	T
27	I	I	I	I	I	I	V	V	V	V	V
67	N	N	N	N	N	N	T	T	T	T	T
74	S	S	S	S	S	S	N	N	V	N	V
83	E	E	E	E	E	E	D	D	D	D	D
91	V	V	V	V	V	V	T	T	T	T	T

Table 3.1.5(B)- Specific nucleotide motifs in different ethnic groups.

Country	China		Korea	Japan	Vietnam	Pacific	USA	Italy		Scotland	
Sub/genotype	adr/C	adw2/B	adr/C	adr/C	ayw1/A	adrq-/C	adw2/A	adw2/A	ayw2/D	adw2/A	ayw2/D
Nucleotide Position											
1934	**T**	**T**	**T**	**T**	**T**	**T**	**A**	**A**	**A**	**A**	**A**
1951	T	T	T	T	T	T	G	G	G	G	G
1979	**A**	**A**	**A**	**A**	**A**	**A**	**G**	**G**	**G**	**G**	**G**
1981	**T**	**T**	**T**	**T**	**T**	**T**	**C**	**C**	**A**	**C**	**A**
2005	T	T	T	T	T	T	A	A	A	A	A
2020	G	G	G	G	G	G	A	A	A	A	A
2059	A	G	A	A	A	A	T	T	T	T	T
2074	T	T	T	T	T	T	C	C	A	C	A
2080	G	G	G	G	G	G	C	C	T	C	T
2092	G	G	G	G	G	G	A	A	A	A	A
2100	**A**	**A**	**A**	**A**	**A**	**A**	**C**	**C**	**C**	**C**	**C**
2104	G	A	G	G	G	G	A	A	A	A	A
2120	**A**	**A**	**A**	**A**	**A**	**A**	**A**	**A**	**G**	**A**	**G**
2121	**G**	**G**	**G**	**G**	**G**	**G**	**A**	**A**	**T**	**A**	**T**
2149	**A**	**A**	**A**	**A**	**A**	**A**	**T**	**T**	**C**	**T**	**C**
2150	T	T	T	T	T	T	C	C	C	C	C
2160	G	G	G	G	G	G	A	A	G	A	G
2171	**G**	**G**	**G**	**G**	**G**	**G**	**A**	**A**	**A**	**A**	**A**

Note: A, specific amino acids, and B, specific nucleotides found in Southeast Asian (red coloured) and Caucasian populations (blue coloured). Amino acids are numbered from the beginning of the HBcAg, and nucleotides numbered from the HBcAg, using the single letters. Bold letters and numbers indicate missense mutations.

Table 3.1.6- Unique HBcAg amino acid and nucleotide changes in specific countries.

Country	Amino Acid Position	Nucleotide Position
India	I80	T_{12}-C_{55}-T_{57}-A_{111}-G_{132}-A_{238}-G_{255}-T_{291}
Pacific	A35-D40-V59-T74-I91	T_{66}-G_{75}-G_{103}-G_{105}-T_{120}-A_{189}-C_{221}-T_{255}
Italy	-	A_{342}
Korea	G87	C_{289}

Note: Amino acids and nucleotides are numbered from the beginning of the HBcAg, using the single letters code.

Table 3.1.7- Correlation of nucleotide sequences of C genes of ayw2 strains between different geographic regions.

Subtype	Region	12	55	57	75	78	103	105	111	120	132
ayw2	India	T	C	T	A	A	T	A	A	A	G
ayw2	Pacific	C	T	G	T	C	G	G	G	T	T
ayw2	Italy	C	T	G	T	A	T	A	G	A	T

Subtype	Region	138	189	207	221	228	238	239	243	246	255
ayw2	India	A	G	C	G	A	A	T	C	G	G
ayw2	Pacific	G	A	T	C	G	G	C	T	G	A
ayw2	Italy	G	G	T	T	A	G	C	T	A	A

Subtype	Region	267	288	291	298	342	346	501
ayw2	India	T	A	T	C	G	A	G
ayw2	Pacific	C	G	C	T	G	C	A
ayw2	Italy	C	G	C	C	A	A	G

Note:Nucleotides are numbered from the beginning of the HBcAg.

Table 3.1.8- Correlation of nucleotide sequence of C-gene of adrq+ and adrq- strains between different genotype C-regions.

Nucleotide Position	China (adrq+)	Japan (adrq+)	Korea (adrq+)	Pacific (adrq-)
66	C	C	C	T
75	T	T	T	G
105	T	T	T	A
255	A	A	A	T
258	A	C	C	C
271	G	G	G	A
289	A	A	C	A

Chapter 3.2

3.2.1 Aims of Study

As mentioned earlier, a pilot study (Chapter 3.1) revealed that specific motifs (aa and nucleotide) were defined for the Pacific regions (both genotypes C and D), but we now intended to study this in greater depth. Unique sequences obtained from the Pacific group at both aa and nucleotide levels (first part of this chapter), led us to more investigation on core gene variability in this area. We selected randomly 58 samples from HBV-infected patients from 4 different islands. HBV isolates from C and D genotypes were identified. Genotype C is predominant in Vanuatu, Fiji and Tonga, while D is the dominant genotype in Kiribati. HBs Ag positive samples were collected from four Pacific islands (Table 3.2.1): Kiribati (14), Vanuatu (14), Fiji (15) and Tonga (15). Geographically, these islands are located as follows: Kiribati in Micronesia; Fiji and Vanuatu in Melanesia; and Tonga in Polynesia (Fig 3.2.1). They were chosen from a previous study of surface variability and correlation with the history of human populations and immigration patterns of the Pacific region (Basuni 2001). Amplifying of C and S genes carried out by PCR, followed by direct sequencing. After construction of simple Neighbour-joining Trees based on core and surface proteins further phylogenetic trees were constructed using sequences from the database. Genotypes C and D strains from the database and from chapter 3.1 (C gene variability and ethnicity) were also included for comparison. Tables showing alignments of amino acid and nucleotide sequences are shown in the appendices at the back of this thesis.

3.2.2 Results

3.2.2A Genotype and subtype prevalence in Pacific islands

Table 3.2.1 shows the distribution of genotypes and subtypes of strains studied from individual islands. The following preparations were genotype D: in Kiribati, 13 out of 14; in Fiji, 5 out of 15; in Vanuatu, 1 out of 14; and in Tonga, 2 out of 15. The remaining sequences

were genotype C. Subtype adrq- was the main subtype in genotype C-occupied islands while ayw2 was the prevalent one in Kiribati.

3.2.2B Genotype/subtype-specific substitutions common to Pacific region

Tables A.I.3 to A.I.14 show the complete aa and nucleotide sequences of all four Pacific island strains (see Index, BioEdit tables). Comparing genotypes C and D, there were 11 aa variations at 10 positions (Table 3.2.2): S/T12, I/V27, A/S35, D/E40, I/V59, N/T67, S/T74, D/E83, I/V/T91 and F/I97. These variants were not island-specific with the exception of V91 found only in Vanuatu.

It was also possible to identify nucleotide variants for genotypes C and D, subtypes adrq- and ayw2 (Tables A.I.12 and A.I.14). Overall, 90 genotype/subtype-specific nucleotide substitutions in 45 positions were found (Table 3.2.3), of which 67 were silent. 23 were missense (genotype specific) in the 10 positions detailed above. Interestingly, a majority of adrq- strains derived from Vanuatu (but not other islands) contained nucleotide variants identical to ayw (genotype D) in 4 positions: 78, 138, 234 and 255.

3.2.2C Substitutions that distinguish specific islands

Island-specific nucleotide variants were identified which increased or decreased in frequency from West to East (from Vanuatu to Fiji and on to Tonga) (Figure 3.2.1, table 3.2.4). Nucleotide changes at positions C_{36}, T_{109}, C_{114}, G_{255} and T_{267} were found in a majority of strains from Tonga. On the other hand, changes including C_{78}, T_{234} A_{246}, A_{255}, G_{271}, C_{354}, C_{366}, A_{454} and G_{510} were predominant in Vanuatu. However, no island-specific nucleotide variants were detected in Fiji and Kiribati. All nucleotide changes that were specific to an island group, except G_{271} (which corresponded to residue V91 in Vanuatu), were silent mutations.

There were six strains from Tonga (3309, 3415, 3417, 3419, 3428 and 3629) and four from Vanuatu (5017, 5022, 5230 and 5072) of genotype C, adrq- subtype, which showed high variability at the nucleotide level compared to other Pacific strains (Tables A.I.8 and A.I.10.

However, in genotype D-studied strains, the only major heterogeneity was seen in isolate 2143 from Kiribati (Table A.I.6). All these strains originated from different cities in each island with various HBeAg/anti-HBe status.

3.2.2D Comparison with international database

Choosing the correct reference sequence is mandatory, especially in cross-sectional studies, otherwise over- and/or under-estimations can occur. The adrq- subtype is rarely found in Genbank; of eight available, six were submitted as part of our previous study (Jazayeri et al, in press) and two isolates (HMA and Cha), published by Norder et al., (7), constituted our reference sequences. No ayw sequence from Pacific (excluding this study) has been reported in Genbank. We chose one genotype C from a Chinese (A35) and one Indian genotype D (0123), both from our pilot study for comparison between corresponding Pacific strains. For genotype C, there were 2 amino acid and 12 nucleotide variations between our predominant Pacific genotype C sequence and A35 (Chinese reference) (Table A.I.11 and A.I.12). However, in genotype D, we found 5 amino acids and 30 nucleotides different between our Pacific D and 0123 (Table A.I.13 and A.I.14).

3.2.2E Phylogenetic Analysis

3.2.2E-I Genotype C Tree

Figure 3.2.2 shows that all the genotype C sequences, from the pilot study (green codes), database (blue codes) and adrq- strains (from this study, red codes), grouped into two major clusters. The first is completely occupied by Pacific adrq-, including one of the database sequences, X75656 (7); a majority of strains belonging to Vanuatu composed a major sub-cluster. The second cluster contains adrq+ sequences (including from the database) derived mainly from S.E Asian populations, which are heterogeneous.

3.2.2.E-II Genotype D Tree

Figure 3.2.3 shows alignment of all genotype D sequences. They grouped into 2 major clusters. The first, similar to the genotype C tree, was entirely made up of Pacific strains (red codes). The second cluster contained two sub-clusters: the first had those from Western populations (a majority from Europe); the second had a combination of Asian and a few European sequences. Inside the Euro-Asian cluster, there was an Indian sub-cluster. Unlike genotype C, the Pacific genotype D shows relative homogeneity with only a few nucleotide mutations, indicating that genotype D arrived recently in the Pacific as a relatively homogenous population. We have been unable to identify a likely source of this migration in this study.

Fig 3.2.1- Pacific map showing boundaries of the Pacific islands and surrounding countries. Fiji and Vanuatu located in Melanesia, Tonga and Kiribati located in Polynesia and Micronesia, respectively.

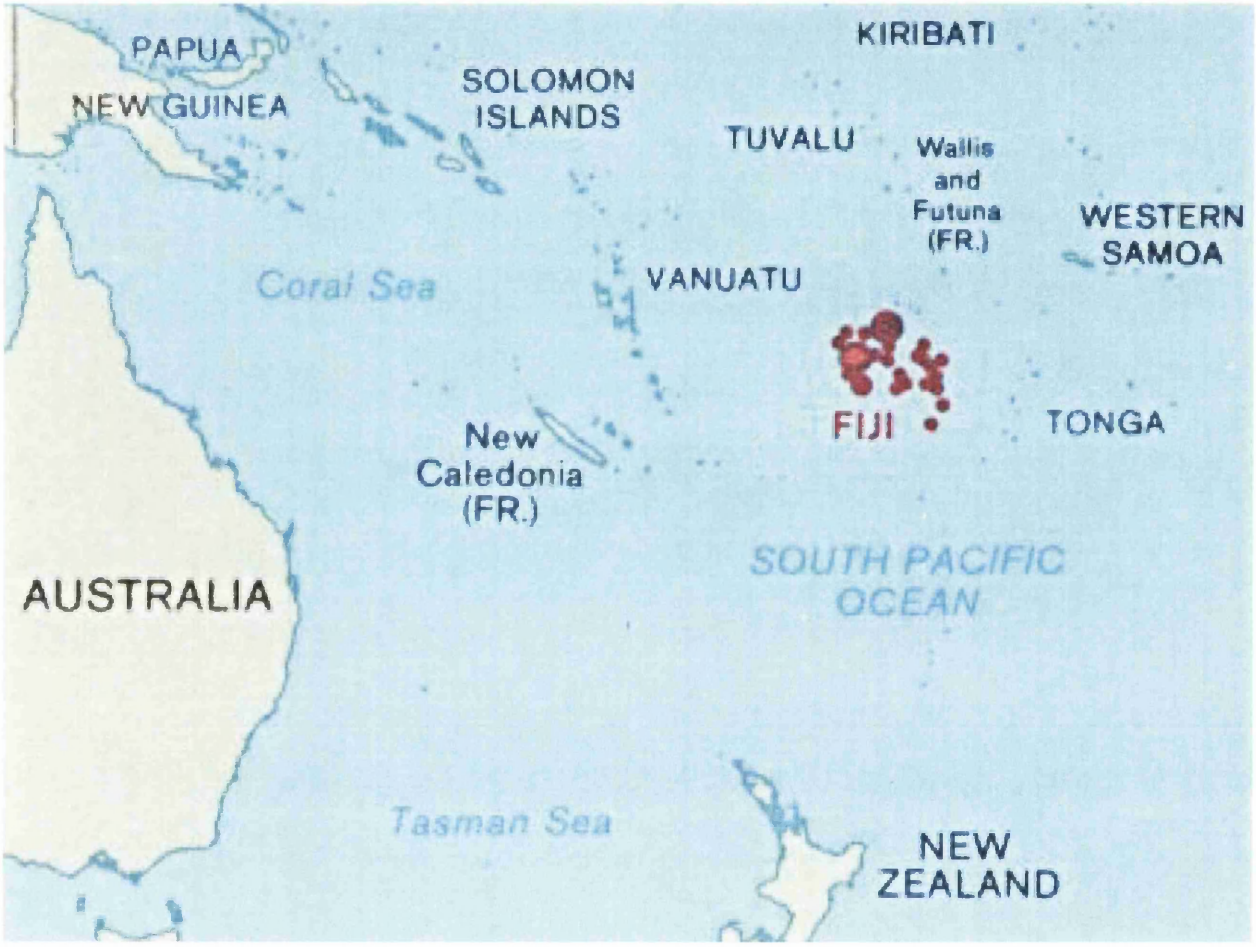


Table 3.2.1- Origin and Distribution of HBsAg positive sera that were used in this study as well as genomic/subgenomic identification of Pacific islands strains.

Island	Sample Code	Genotype	Subtype	Island	Sample Code	Genotype	Subtype
Fiji				**Vanuatu**			
	0142	C	adrq-		5022	C	adrq-
	0143	C	adrq-		5072	C	adrq-
	0314	C	adrq-		5073	C	adrq-
	0349	C	adrq-		5075	C	adrq-
	0480	C	adrq-		5081	C	adrq-
	0550	C	adrq-		5093	C	adrq-
	0800	C	adrq-		5114	C	adrq-
	0880	C	adrq-		5311	C	adrq-
	0026	C	adrq-		5186	C	adrq-
	0075	C	adrq-		5230	C	adrq-
	0078	D	ayw2		5264	C	adrq-
	0079	D	ayw2		5265	C	adrq-
	0092	D	ayw2		5017	C	adrq-
	0456	D	ayw2		5029	D	ayw2
	0470	D	ayw2				
Tonga				**Kiribati**			
	3415	C	adrq-		2080	C	adrq-
	3417	C	adrq-		2006	D	ayw2
	3097	C	adrq-		2007	D	ayw2
	3419	C	adrq-		2019	D	ayw2
	3428	C	adrq-		2119	D	ayw2
	3629	C	adrq-		2127	D	ayw2
	3099	C	adrq-		2317	D	ayw2
	3221	C	adrq-		2483	D	ayw2
	3365	C	adrq-		2109	D	ayw2
	3369	C	adrq-		2110	D	ayw2
	3519	C	adrq-		2117	D	ayw2
	3309	C	adrq-		2143	D	ayw2

3509	C	adrq-	2039	D	ayw2
3414	D	ayw2	2084	D	ayw2
3343	D	ayw2			

Note: 10 of the above samples had been used in the pilot study (see tables A.I.1 andA.I.2).

Table 3.2.2. The most frequent differences in core gene amino acid sequences between genotypes/subtypes of the Pacific region.

Amino acid Position	Genotype D Subtype ayw	Genotype C Subtype adrq-
12	T	S
27	V	I
35	A	S
40	D	E
59	I	V
67	T	N
74	T	S
83	D	E
91	T	I/V*
97	F	I

Note: Each specific amino acid was found in a majority of strains from each genotype/subtype and the combination was unique for that genotype/subtype. Amino acids are represented by single letters and are numbered from the beginning of the core gene.

V* was only seen in adrq- strains from Vanuatu.

Table 3.2.3. Unique C-gene nucleotide differences which permit genotypes C and D classification.

Nucleotide Position	Genotype D Subtype ayw	Genotype C Subtype adrq-	Nucleotide Position	Genotype D Subtype ayw	Genotype C Subtype adrq-
34	**A**	**T**	**200**	**C**	**A**
51	G	T	204	A	G
66	C	T	219	T	A
75	T	G	**221**	**C**	**G**
78*	C	T	228	A	G
79	**G**	**A**	234‡	T	C
81	**A**	**-**	243	T	C
90	T	C	**249**	**C**	**A**
93	A	C	250	C	T
96	T	C	255†	A	G/T
103	**G**	**T**	**272**	**C**	**T**
105	**G**	**A**	288	G	A
120	**T**	**G**	**289**	**T**	**A**
135	T	G	308	T	C
138Φ	G	A	324	C	T
159	T	A	342	G	T
174	A	T	348	A	T
175	**A**	**G**	387	C	T
180	T	G	390	A	C
183	C	T	396	T	C
189	A	T	418	C	T
192	A	G	435	G	A
193	C	T			

Note: Nucleotides are represented by single letters and numbered from the beginning of the core gene. Dash lines represent no variant specified for that position. Bold numbers and letters indicate amino acid altered positions (missense mutations).

***: A majority of adrq- strains from Vanuatu had C_{78}.**

$^{\Phi}$: A majority of adrq- strains from Vanuatu had C_{138}.

‡: adrq- strains from Vanuatu contained T_{234}.

†: A_{255} seen in a majority of samples from Vanuatu despite belonging to adrq- subtype.

Table 3.2.4. Nucleotide differences between prevalent strains in Fiji, Tonga and Vanuatu.

Nucleotide Position/changes	Vanuatu	Fiji	Tonga
C36T	2	2	9
C78T	10	1	1
T109C	1	2	9
C114T	0	4	15
G138A	11	1	1
T234C	11	0	1
A246G	11	1	1
G255T	0	2	7
A255T	9	1	0
T267C	0	2	9
G271A	12	1	0
C354T	11	2	0
C366T	11	1	0
A454C	7	2	0
A501G	11	5	0
G510A	11	2	0

Fig 3.2.2. Neighbour joining phylogenetic tree of core gene sequences of genotype C from Pacific adrq- strains (red colour), adr subtypes from the chapter 3.1 (green colour) and database (blue colour) strains rooted with samples I98. Figure shows bootstrap value of 70% and scale denotes percent diversity.

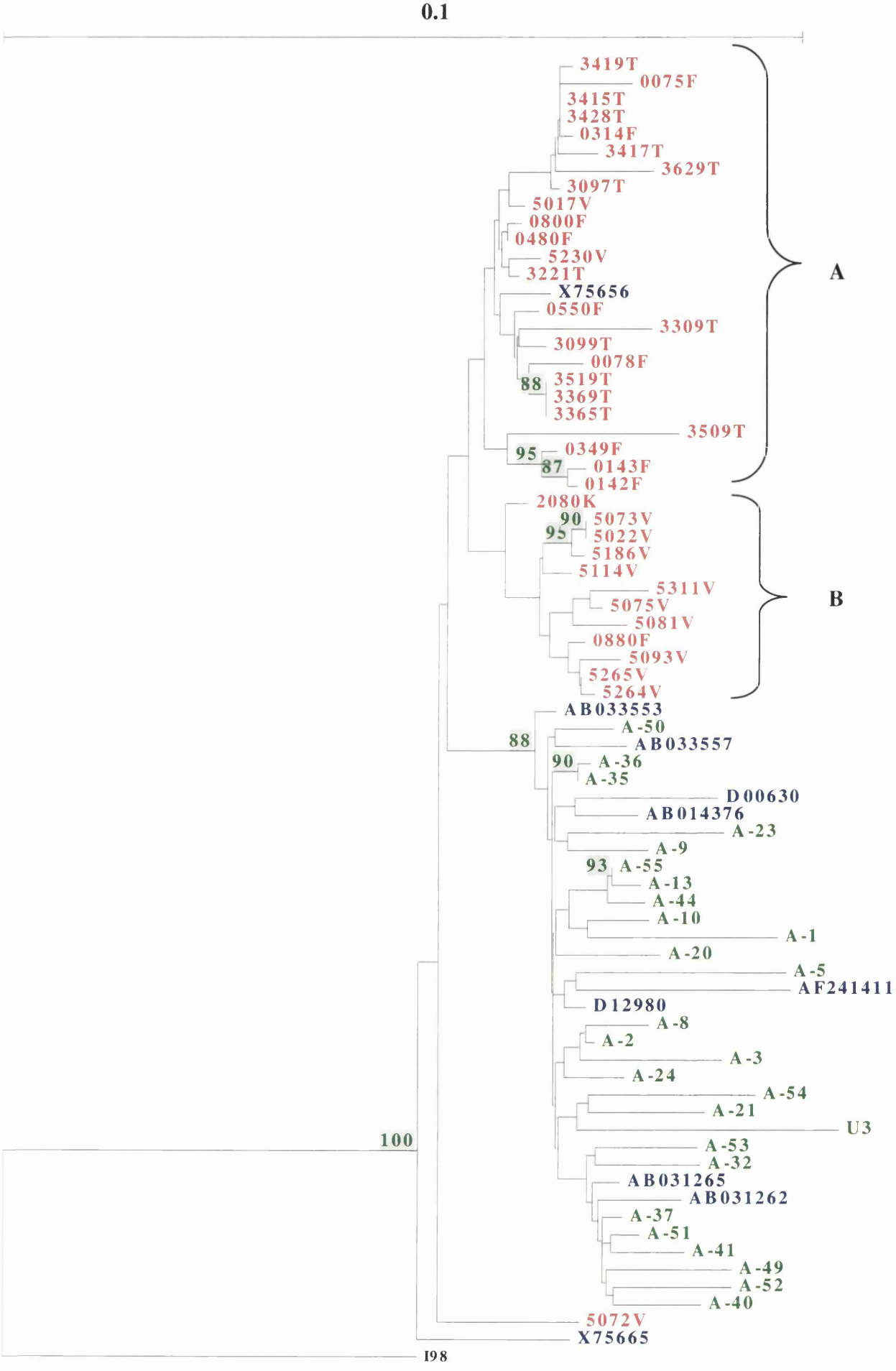
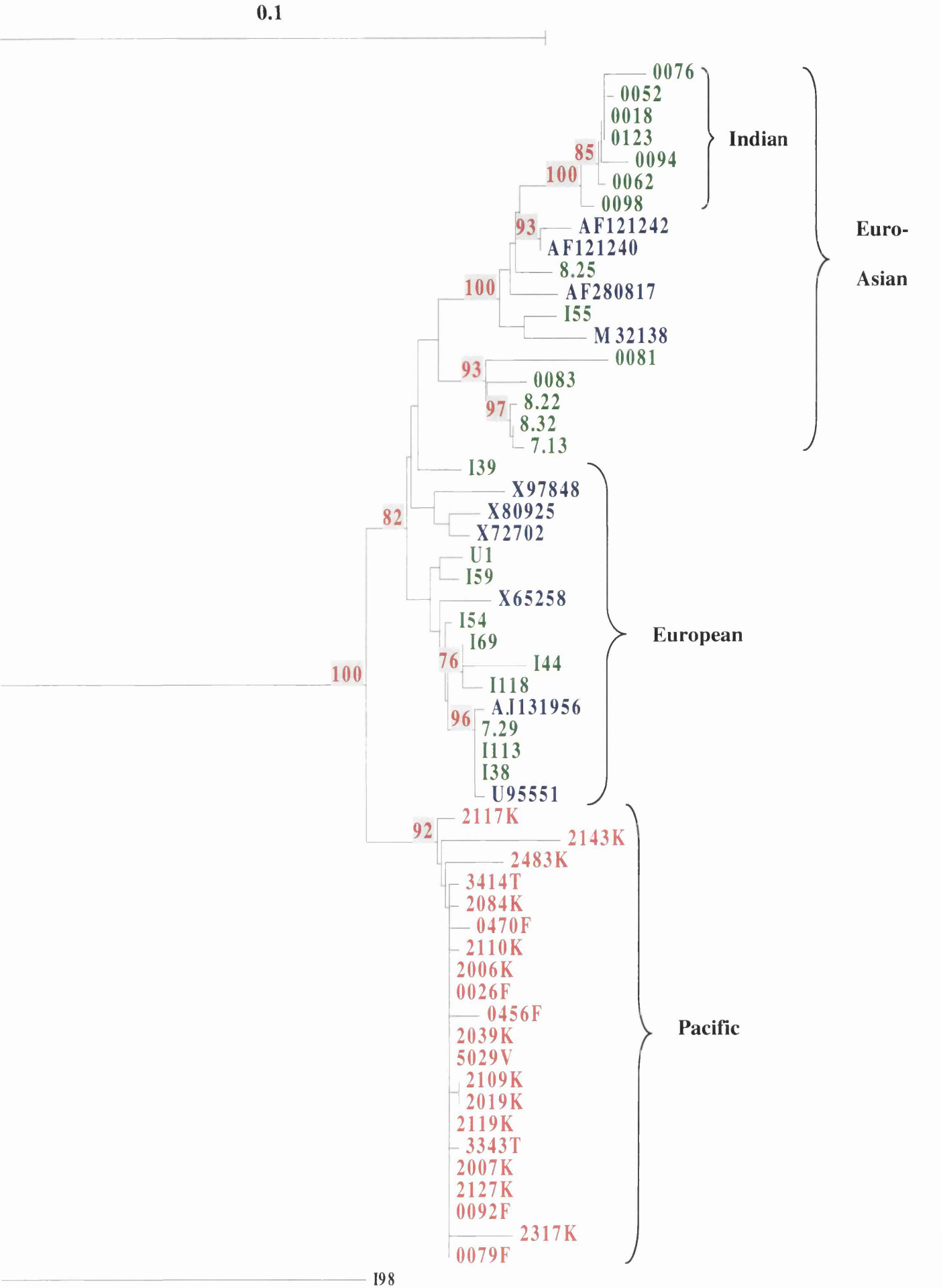

0.1
3419T
0075F
3415T
3428T
0314F
3417T
3629T
3097T
5017V
0800F
0480F
5230V
3221T
X75656
0550F
3309T
3099T
0078F
88
3519T
3369T
3365T
3509T
95
0349F
87
0143F
0142F
A
2080K
90
5073V
95
5022V
5186V
5114V
5311V
5075V
5081V
0880F
5093V
5265V
5264V
B
AB033553
A-50
88
AB033557
90
A-36
A-35
D00630
AB014376
A-23
A-9
93
A-55
A-13
A-44
A-10
A-1
A-20
A-5
AF241411
D12980
A-8
A-2
A-3
A-24
A-54
A-21
U3
100
A-53
A-32
AB031265
AB031262
A-37
A-51
A-41
A-49
A-52
A-40
5072V
X75665
I98

Fig 3.2.3. Neighbour joining phylogenetic tree of core gene sequences of genotype D from the Pacific (red colour) and chapter 3.1 (green colour) and database (blue colour) ayw strains, rooted with samples I98. A majority of Indian sequences which form a separate sub-cluster in the Euro-Asian set, are indicated. Figure shows bootstrap value of 70% and scale denotes percent diversity.


0.1
0076
0052
0018
0123
85
0094
100
0062
0098
Indian
93
AF121242
AF121240
8.25
100
AF280817
I55
M32138
0081
93
0083
97
8.22
8.32
7.13
Euro-
Asian
I39
X97848
82
X80925
X72702
U1
I59
X65258
I54
I69
76
I44
100
I118
A.I131956
96
7.29
I113
I38
U95551
European
2117K
92
2143K
2483K
3414T
2084K
0470F
2110K
2006K
0026F
0456F
2039K
5029V
2109K
2019K
2119K
3343T
2007K
2127K
0092F
2317K
0079F
Pacific
I98

3.3 Discussion

The prevalence of hepatitis B in South East Asia and western Pacific (except in Australia, New Zealand and Japan, where the mean carrier rate is less than 2%, see chapter 1.4) is amongst the highest in the world; more than 75% of the world's chronic carriers are living in such densely populated regions (Maynard et al., 1989). Since the world-wide distribution of HBV follows a geographic pattern (Norder et al., 1993b), there may be a strong influence of ethnic background, perhaps driven by T-cell selection, on this distribution, reflected by divergence of amino acid substitutions within certain regions of the core gene (Bertoletti et al., 1991; Zampino et al., 2002).

HBV antigenic subtypes generally reflect the country or regions of origin (Couroucé-Pauty et al., 1983). However, there are a number of confounding factors. First, some subtypes are not geographically localized but are worldwide. Second, there are differences in HBV strains between native and foreign carriers in each geographic area (Couroucé-Pauty et al., 1983); this means that genetically homogenous HBV strains cannot simply be presumed to be linked epidemiologically. Third, some subtypes are genetically heterogeneous and belong to more than one genotype (see chapter 1.5, table 1.2) (Norder et al., 1992). Knowing the degree of HBV variation to be expected in a certain community could be useful in epidemiological investigations (Bläckberg et al., 2000).

Our comparison of 139 new complete core genes and those in databases verified that amino acid/nucleotide specific substitutions correlate with both ethnicity and HBV genotypes/subtypes. Motifs and specific individual mutations were observed that correlated with broad ethnic background, country, genotype and subtype. Our finding that HBcAg variability was distributed geographically was entirely consistent with Norder's study of S genes (Norder et al., 1993b). Moreover, construction of phylogenetic trees (Figure 3.1.1), of C and S genes showed an almost identical pattern of HBV genotype

distribution. Some samples had discrepant genotype allocation based on analysis of the C and S genes. These discrepancies were not surprising, as recombination events between different genotypes have been described previously (Mizokami et al., 1997); between A and C from Vietnam (Hannoun et al., 2000b), between C and D from Tibet (Cui et al., 2002) and between B and C from different countries of South-East Asia (with exception of Japan) (Suguachi et al., 2002). Of interest is that sample A-5 came from Korea, where recombination between genotypes B and C has been reported, designed as genotype Ba (a for Asia). This is in keeping with Suguachi's study (Suguachi et al., 2002). In our study, recombination between A and D was found in two Italian sequences (Figure 3.1.1). Below, we briefly discuss some aspects of our findings.

Comparisons of our data with HBcAg sequences in databases (except ayw1 strains, of which there were no examples) were almost entirely consistent. In other's work, codon 97 showed the highest variation (Yuan et al., 1999, 2000); however, here, codon 74 contained the most discrepancy between sequences. Of 9 variable residues observed throughout the C gene, 7 were characteristic for genotypes A and D, and 2 residues were characteristic for the adrq- subtype. As some subtypes belong to more than one genotype, these variants may also be found in more than one genotype. Further variants were identical in specific subtypes allowing proper categorization.

The findings at the amino acid level were mirrored at the nucleotide level. At the nucleotide level (Table 3.1.4), the best correlation between our data and those obtained from databases was found in genotypes A (adw2 subtype) and D strains. Because 6 of 8 adw2 (genotype B) sequences in our study were derived from Chinese patients, we compared these to other Chinese adw sequences in the database (Chuang et al., 1993). These were almost identical. Adrq+ strains in the databases were also very similar to ours, but adrq- strains did not contain T_{75}. Tables 3.1.3 and 3.1.4 propose amino acid and

nucleotide substitutions that are critical sites for defining genotypes and their corresponding subtypes.

It is particularly interesting that peoples of a broad ethnic background share amino acid substitutions in the core gene despite having different subtypes (Table 3.1.5A). This was particularly true of samples from the South-east Asian (including Pacific) regions. It would appear that there are shared host-factors, most likely immunological, which act on these sequences despite their background subtype. Of interest is that Asian immigrants to the USA shared identical sequences to those from the home country. In keeping, in a large -scale study on HBV genotype distribution in the USA between different ethnic groups, Chu et al, (2003b) found that among the patients born outside the United States, the distribution of HBV genotypes was related to the prevalent HBV genotype in the place of birth, and between all epidemiologic factors, only ethnicity and place of birth were independently related to HBV genotypes. It thus appears that there are no signature sequences in the USA, probably because of long-term, widespread importation of sequences. One of our hypotheses was that specific ethnic groups would maintain the virus that originated in their home country. Immigration to the USA had no effect on Asian sequences, in keeping with circulation of HBV amongst ethnic minorities or the likelihood that in an intra-ethnic sexual partnership, it is the Asian partner who is more likely to be the primary infecting source (because of high prevalence rate of infection, see chapter 1.5). Nevertheless, genotype/subtype-specific variants were identical whatever the geographic origin.

The pattern in Western-derived sequences (including from US-Caucasians, Scotland and Italy) was slightly different, as there were no specific substitution that correlated with individual countries (Table A.I.2). Nevertheless, a clear motif was identified that was common to Western populations. We did not study genotype F, which is native to the Americas.

The Pacific region comprises a unique region of admixing of two HBV genotypes, C and D, with specific identities which have not been identified in other regions. Adrq+ is the prevalent subtype in South-East Asia, while adrq- is only found in Oceania. Geographical transition from Asian adrq+ to South American adw4q- has been suggested to be via the adrq- subtype in Oceania (Couroucé-Pauty et al., 1983). The gradient of nucleotide and aa variations from west to east in our study (shared sequences between Tonga and Fiji and, on the other hand, variations between Tonga and Vanuatu), are most consistent with the hypothesis of migration of Polynesian people from Southern China through Melanesia and Fiji and their radiation across the Pacific to fill the Polynesian triangle in different times (Diamond et al., 1988; Hagelberg and Clegg, 1993; Redd et al., 1995; Kayser et al., 2000). In contrast, back migration from Polynesia to Melanesia could also be responsible for this variation gradient (see below). Of interest is that 10 strains from Vanuatu, all genotype C isolates, constitute a subcluster in genotype C tree (Figure 3.2.2, cluster B). However, two isolates from Vanuatu were extremely divergent and found within subcluster A. Within the subcluster A, 6 strains from Tonga constitute a separate branch (see Figure 3.2.2). The presence of different strains from all islands (in particular Fiji and Tonga) in cluster A may indicate admixture of isolates in Fiji due to its geographical location between other 3 islands (Figure 3.2.1). The gradient of nucleotide substitutions between 3 islands in table 3.2.4 showed commonality between Fiji and Tonga, rather than Vanuatu. This is in keeping with the history of an isolated Vanuatu.

In contrast, genotype D strains including those from databases showed that the Pacific islands contain a unique homogenous cluster separated from other genotype D groups. The few D sequences on other islands show a similar homogeneity, indicating recent contact between Kiribati (which is almost an isolated island in Micronesia with limited transfer of virus-see figure 3.2.1) and the genotype C-dominated islands. This suggests

that genotype D arrived in the Pacific as a relatively homogenous population and has been distributed by admixing with stable human populations (Basuni, unpublished data). As Indian immigrants usually have genotype D, one might expect Indians to be the origin of genotype D strains in this region. However, our results did not show such an acquisition of these new strains by migration, there was substantial variability between these two ethnic groups. This is best seen as the two separate clusters in figure 3.2.3. The study on S genes by Basuni et al, also revealed no strong evidence for Indian admixture in Pacific genotype D populations. In fact, results from our previous pilot study showed that at the nucleotide level the Pacific genotype D sequences were closer to Caucasian strains rather than Indian (Tables A.I.1 and A.I.2). Thus, a Pacific origin for genotype D cannot be ruled out.

In the Pacific, the contrast between the silent mutations that predominate within Pacific genotype C sequences and non-synonymous changes that separate Pacific from the S.E Asian sequences is suggestive of an interaction with human genetic variants, especially HLA, in the evolution of geographically separated HBV lineages. Once in the Pacific, no further immune mediated amino acid selection seems to have occurred, probably because there are only a few Pacific HLA types. On the other hand, the significant nucleotide variability is in keeping with random mutation in isolated island population.

Overall, the high diversity for genotype C (including silent mutations) and the island-specific sequences, suggest a long history of evolution and isolation for genotype C-islands. In contrast, the low diversity of genotype D indicates a recent, limited, spread over the rest of the Pacific with a single lineage.

HBcAg contains helper T cell (Jung et al., 1995) and cytotoxic T lymphocyte (CTL) epitopes (Ehata et al., 1992; Chuang et al., 1993; Wakita et al., 1991). Different HLA restricted CTL/T helper epitopes have been identified within the core gene which are

capable of inducing significant T-cell responses in HBV infected patients. (Ferrari et al., 1991; Ehata et al., 1992,1993). Previous results have suggested that epitopes for CTL/T h recognition might be different on account of the diverse distribution of HLA antigens in different geographic regions (Thursz et al., 1995; Thio et al., 1999). Core variation was most likely to occur at positions within known epitopes which can tolerate naturally occurring variants (Bertoletti et al., 1991; Ehata et al., 1994; Akarca et al., 1995; Hosono et al., 1995; Ferrari et al., 1996; Yuan et al., 1999, 2000) i.e. residues 12, 27, 67, 74, 83, 87, 91, 97 and 116 (Table A.I.1). In keeping, a majority of nucleotide variants, either country-specific or subtype/genotype-specific, were distributed in known dominant epitopes including regardless whether silent or missense mutations, residues 27, CTL (Bertoletti et al., 1991); 50-69, Th (Ferarri et al., 1991); 74-89 and 107-118, B cell epitopes (Colucci et al., 1988; Salfeld et al., 1989). Most changes were synonymous. This might be due to a lack of a positive selection advantage or a strong bias towards maintaining that amino acid. The similarity in amino acid/nucleotide distribution in adw2 (genotype A) and ayw2 (genotype D) strains found in Western countries (USA-Caucasian, Italy, and Scotland), showed that considerable constraints must exist against HBV variability in a particular genotype infecting a person of a particular background. In addition, the "South-East Asian" and "Western" motifs suggests either that there is positive selection on specific variants or that genetic drift in HBV is relatively slow (Wilson et al., 1999). According to the almost unique amino acid variation in residues 35, 40, 59, 74, 80, 87 and 91 in particular countries (Table 3.1.6), we hypothesize that differences in distribution of HLA antigens, or other immune genes between diverse geographic areas (Ehata et al., 1994) (see chapter 1.7.1), probably contributed to the selection of amino acid variation. However, primary data need to be accumulated to investigate this issue.

Chapter 4 In Vitro HBcAg Localisation

Subjects and study design

The third section of this thesis deals with in vitro biological effects of HBV C gene variability. It has been shown that, after seroconversion to anti-HBe, multiple amino acid substitutions appear in different regions of the hepatitis B core protein. It is not known if these have any biological significance and little is known about the cellular localisation of HBcAg variants. If sequence could be shown to correlate with distribution of HBcAg, especially if the distribution changes with selection of mutations in individual patients, this may give important clues to potential mechanisms of liver damage. To address this issue, confocal microscopy was used to visualise HBcAg cellular expression in 40 clones (including mutagenised samples) from various stages of HB disease from patients with different HBeAg/anti-HBe status in the presence of cell cycle blocking agents. This was correlated with sequence variation within B cell epitopes and the C-terminus. Then, using mutagenesis, we directly investigated the influence of these mutations on intracellular trafficking.

14 Greek patients were selected after previous characterisation of their sequential HBV pre-core/core sequences with known serological markers of HBV infection (Carman et al., 1996). 26 clones were studied. They fell into three groups: six patients, had a sample taken from the time when HBeAg was positive (Table 4.1.1 group1-A, including 2 initial samples taken from the first patient, $1A^1$ and $1A^2$), plus a second sample after anti-HBe seroconversion (Table 4.1.1 group 1-B); five patients, were from continuously anti-HBe positive with active disease, in whom one sample was chosen from an early period when few HBcAg mutations were apparent (Table 4.1.1 group 2-A), and a second after selection of either a pre-core stop codon and/or at

least one HBcAg mutation (Table 4.1.1 group 2-B). Three HBeAg-positive chronic patients provided single time-point samples (Table 4.1.1 group 3-A).

PCR was employed to detect and amplify the core gene, followed by direct sequencing of the amplicon. Cleaned PCR products were digested with restriction enzymes, then ligated into a eukaryotic expression vector (pRK5), transformed with competent E.coli DH5α, and a large scale DNA preparation was obtained and then sequenced. After transfection of the cloned samples with a liposome-based method in the presence of cell cycle blocking agents (confirmed by flow cytometry), localisation of core protein was studied using immunofluorescence and confocal microscopy. Correlation between HBcAg localisation and core gene sequences was then analysed. Next, certain samples were selected for a PCR-based site-directed mutagenesis. This was carried out using oligonucleotides to revert mutated sequences back to the original sequence. This was followed by studies of the resultant intracellular distribution.

4.1 Results

4.1.1 Synchronisation of Cos-7 cells by Thymidine-Aphidicolin

To study the influence of cell cycle on the subcellular distribution of HBcAg, Cos-7 cells were synchronised by two blocking agents (see chapter 2, figure 2.2). In initial experiments (data not shown) aphidicolin used alone soon after transfection yielded low transfection efficiencies. Consequently, a two step cell cycle synchronisation was employed. As a first step, cells were incubated in the presence of thymidine for 12 hours to enter G0/G1 phase of the cell cycle (Figure 4.1.1, panel B). Afterwards, cells were withdrawn from thymidine and permitted to progress out of G0/G1 and into the S, G2, and then M phases of the cell cycle for 12 hours (Figure 4.1.1, panel C). After cell cycle release, they were washed twice with PBS and transfection was carried out

Table 4.1.1- Identification, serologic data and clinical status of HBsAg positive sera that were used as the source for HBcAg cellular expression with different serological and clinical status.

Sample Code	Genotype	HBeAg/ anti-HBe	Clinical Outcome	HBcAg Localisation	B-cell Epitope Mutation	C-terminus Mutation	Pre-core Stop Codon Mutation
Group 1							
1A[1]	A	+/-		N	-	-	-
1A[2]		+/-		C	+	+	-
1B		-/+	Remission	C	+	+	-
2A	D	+/-		N	-	-	-
2B		-/+	Remission	C	-	+	+
3A	D	+/-		N	-	-	+
3B		-/+	Remission	C	+	+	+
4A	D	+/-		N	-	-	-
4B		-/+	Active	C	+	+	+
5A	D	+/-		N	+	-	-
5B		-/+	Active	C	+	+	+
6A	D	+/-		N	+	+	+
6B		-/+	Active	C	+	+	+
Group 2							
7A	D	-/+		NC	+	+	-
7B		-/+	Remission	NC	+	+	+
8A	D	-/+		N	-	+	+
8B		-/+	Active	C	+	+	+
9A	D	-/+		N	+	-	-
9B		-/+	Active	NC	+	+	+
10A	D	-/+		C	+	-	+
10B		-/+	Active	N	-	-	-
11A	D	-/+		N	+	+	+
11B		-/+	Active	C	+	+	+
Group 3							
12A	D	+/-	Chronic disease	N	-	-	-
13A	A	+/-	Chronic disease	N	-	-	-
14A	D	+/-	Chronic disease	N	-	-	-

Note: These are three groups of patients: HBeAg to anti-HBe seroconverters, group 1; continuously anti-HBe positive, group 2, and single samples from HBeAg positive patients, group 3. There are two early samples taken from patient 1 ($1A^1$ and $1A^2$). Pre-core stop codon mutation is A_{1896}. Distribution of HBcAg is as follow: N, nuclear; C, cytoplasmic; NC, nuclear and cytoplasmic. +, presence of mutation; -, absence of mutation. Note that the pre-C region was not within the cloned product so did not directly affect localisation.

using FuGENE-6 and/or Lipofectamin Plus Transfection Reagent methods. At the second step, after 12 hours post-transfection, cells were treated with 2.5 μg/ml aphidicolin for 30 hours. Synchronisation by aphidicolin resulted in 94 % of cells remaining in G1/S phase (Figure 4.1.1, panels D and E).

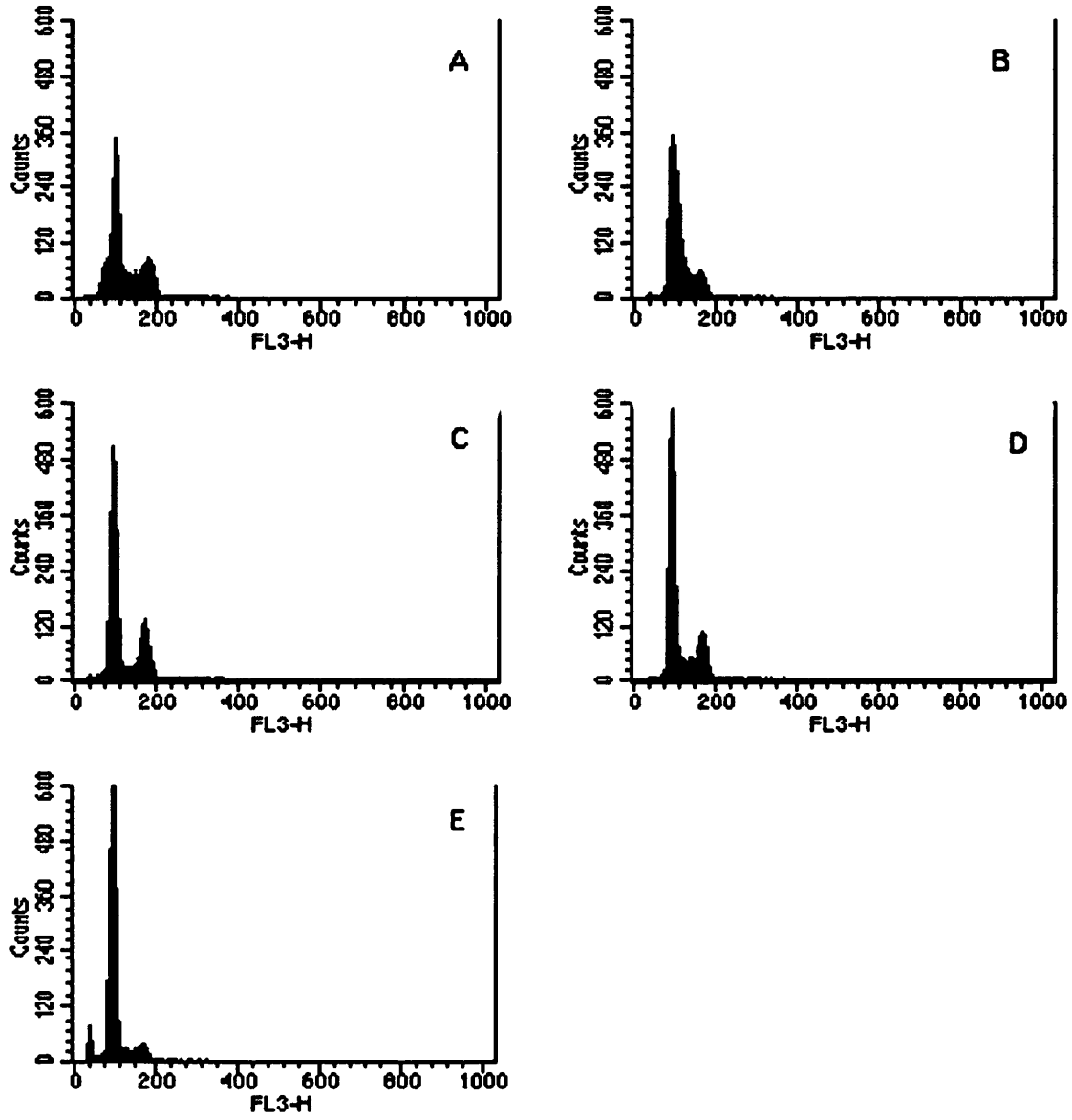
4.1.2 Distribution of mutations in the core gene B cell epitopes and C-terminus region

Overall, 34 mutations were distributed in different known B cell epitope regions within the core gene (see table 4.1.2). 22 (64%) occurred in residues 74-89; 3 (9%) in 107-118; and 9 (26%) in 128-135 (Table 4.1.2). 8 out of 22 mutations in the first B cell epitope, aa residues 74-89, occurred at aa position 80. 13 (59%) out of 22 C-terminus mutations occurred within aa residues 148-160 (Table 4.1.2).

4.1.3 Cellular distribution of core protein

The expression of HBcAg was predominantly nuclear in 13 cases, both nuclear and cytoplasmic in 3 samples, and predominantly cytoplasmic in 10 cases. Of 13 patients who showed predominantly nuclear, 9 were HBeAg and 4 anti-HBe positive (note that 3 of them were initial samples: 8A, 9A, 11A and 10B) (see below). Only one patient with predominantly cytoplasmic distribution was HBeAg positive ($1A^2$), i.e., 9 were anti-HBe positive. In keeping, none with both nuclear and cytoplasmic distribution were HBeAg positive, i.e., all were anti-HBe positive. Also, we observed a form of nuclear expression associated with the membrane using standard microscopy (Figure 4.1.2, 7A and 9B; Figure 4.1.3 3B, 4B, with double mutagenesis, and 11B). Figures 4.1.2 and 4.1.3 and tables 4.1.1 and 4.1.2 detail the patterns of core localisation from all patient-derived HBV core genes cloned into pRK5 and transfected into G0/1-arrested Cos-7 cells. Consequently, our first finding is that cytoplasmic distribution is related closely to HBeAg negativity/anti-HBe positivity.

Fig 4.1.1- FACS anlysis of Cos-7 cells determined by flow cytometry.



Note: The peak to the left corresponds to cells in the G1 phase of the cell cycle; the peak to the right corresponds to cells in the G2 phase. The shoulder between the two peaks corresponds to the S phase. Differential cell cycles shown variable features according to the treatment described in the materials and methods. Asynchronously growing cells were given a characteristic trimodel profile, as negative control (A); the thymidine-treated cells (B) with 84% of cells were in G0/1. The release period in which cells were thymidine-free (C); the aphidicolin-treated cells after 12 hours (D); and after 30 hours (E) post-aphidicolin. 88% and 94% of Cos-7 cells arrested in G1-S phase of the cell cycle after 12 and 30 hours incubations with aphidicolin, respectively.

4.1.4 Core gene sequences from the initial samples showed nuclear distribution

Tables 4.1.1 and 4.1.2 show the results of core gene substitutions in 11 initial samples (either HBeAg or anti-HBe positive) and 3 single time-point HBeAg positive samples. 12 of these 15 samples were predominantly nuclear (only one of the later samples, 10B, had nuclear localisation; in total 13 were predominantly nuclear). 8 out of 13 (all but one were HBeAg positive) did not have any mutation. The remaining 5 (3 of 5 were anti-HBe positive) had mutations in either C-terminal region and/or B cell epitopes. Two of these were mutated only in B cell epitopes: A74V/G (two possible variants for genotype D-see chapter 3.1C, table 3.1.4) and T80A in samples 9A and 5A, respectively. Another two contained mutations in both regions: T80A, Q130P, C153G and T155S in 6A; S135P, Q151R and T155S in 11A. Only one sample with a C-terminus mutation alone was observed: C153G in 8A.

As noted below, later samples from all these five developed further mutations and shifted to the cytoplasm. Finally, the HBcAg from the remaining 2 initial samples, 10A and 7A, showed either predominantly cytoplasmic or both nuclear and cytoplasmic distribution, respectively (Figure 4.1.2). Thus, all initial samples, but two, showed predominantly nuclear expression. It was unusual to observe C-terminus and/or B cell epitope mutations and nuclear localisation of samples with HBeAg positivity.

4.1.5 Accumulation of B cell epitope/C-terminus mutations and HBcAg localisation

In the five patients with initial samples which contained B cell epitope and/or C-terminus mutations and predominantly nuclear localisation (Table 4.1.1: 5A, 6A, 8A, 9A, 11A), there was a shift over time both in cellular distribution and sequence. Later samples (table 4.1.1: 5B, 6B, 8B, 9B, and 11B) had predominantly cytoplasmic

distribution (see below): all sequences in the latter group contained additional mutations in the same and/or other regions (Table 4.1.2). Thus, it appears that there is a relationship between accumulation of further mutations and shift to the cytoplasm.

4.1.6 A shift in cellular distribution to the cytoplasm is linked to substitutions in the carboxy terminal and/or B-cell epitopes

In the later samples, of ten with predominantly cytoplasmic distribution, 8 contained both C-terminus and B-cell epitope mutations (Tables 4.1.2 and 4.1.3). 2B only contained a C-terminus mutation, A160T, while 10A only had mutations in two B cell epitopes: A74V/G and P131A. 10B although taken in the anti-HBe positive phase, shifted back to the nucleus: in keeping, all the mutations in the first sample (10A), reverted back to the wild type (no mutations) in sample 10B (Figure 4.1.2; Table 4.1.2). 8A from the initial phase, which had a C-terminus mutation, C153G (see above), reverted back and L156P arose in the anti-HBe positive phase (Table 4.1.2 sample 8B). In this patient, an additional amino acid variation occurred in the later sample, 8B: G74V. Both of these variants in amino acid position 74 are specific for genotype D (see section 3.1.C; Table 3.1.4). Of two samples which showed both nuclear and cytoplasmic distribution (7B and 9B), both contained mutations in both C-terminus and B cell epitope regions (Table 4.1.2). In sequences with mutations, there were no differences in the distribution of those mutations between the two clinical groups i.e., HBeAg to anti-HBe seroconverters and the continuously anti-HBe positive. Thus, all later samples showed predominantly cytoplasmic (including both nuclear and cytoplasmic distribution) and contained B cell epitope and/or C-terminus mutations. All but one were anti-HBe positive. As a result, cytoplasmic distribution of HBcAg was strongly correlated with the presence of mutations in B cell epitopes and/or the C-terminus region of the C gene and anti-HBe positivity.

4.1.7 Correlation of HBcAg localisation with clinical outcome

Of the 11 later samples, 4 were in clinical remission and seven had active disease. 3 of patients in clinical remission showed a shift in distribution of HBcAg from nucleus to cytoplasm after seroconversion to anti-HBe. Of the seven patients with active disease all but one (10B) showed a shift from nuclear to cytoplasmic expression between the earlier and later sample (Table 4.1.1). Interestingly, with exception of 10B, all active and all remission phase samples contained B cell epitope and/or C-terminus mutations (Table 4.1.2). Consequently, 9 out of 11 later samples either in remission or in active disease shifted from predominantly nuclear to predominantly cytoplasmic indicating that there was no clear difference between different clinical outcomes and HBcAg localisation.

4.1.8 The cellular distribution of the core protein is dependent on the sequence

Some of the shifter samples (Table 4.1.1) were chosen for site-directed mutagenesis (Table 4.1.3). Samples 3B, 4B and 5B had double mutagenesis at the same time on both C-terminus and B cell epitopes. As sample 10A did not contain a C-terminus mutation, but had mutations in two different B cell epitopes (Table 4.1.2), it was selected for double mutagenesis on these epitopes. All experiments were designed to revert the mutated amino acid to the expected aa variant for each genotype which was also that observed in the first sample of the pair (or the second in the case of patient 10). All clones were sequenced to confirm reversion (Figure 4.1.4). After reverting the sequence in the C-terminal and/or B-cell epitopes of the later sample back to the original, a shift back to the original nuclear and/or both nuclear and cytoplasmic distribution was observed as follows. After mutating single amino acids in B cell epitopes, 4 out of 5 samples shifted back partially i.e., they localised to both nucleus and cytoplasm (Figure 4.1.3, 3B, 4B, 10B and 11B). In contrast, mutagenesis of C-

terminus mutations resulted in predominantly nuclear localisation in all cases (Figure 4.1.3, 2B. 4B and 5B; Table 4.1.3). In keeping, all double mutated samples shifted back to either nucleus (Figure 4.1.3, 5B and 10A) or nucleus and cytoplasm (Figure 4.1.3, 3B and 4B). Consequently, reversion of C-terminus mutations leads to HBcAg shifting back to the nucleus. Reversion of B cell epitope mutations leads to a mixed nuclear and cytoplasmic distribution. The effect of dual mutations is the same as for C-terminus mutations.

4.1.9 Pre-core variants

8 out of 13 samples with predominantly cytoplasmic localisation contained the precore stop codon mutation at aa 28 of the pre-core region; in predominantly nuclear cases it was found in 4 out of 13 (Table 4.1.1). Consequently, the pre core region did not appear to influence the effect of C gene localisation.

Table 4.1.2. Laboratory data and results of HBcAg localization according to samples taken from each patient.

Patient Code	*HBc Ag* Loc.	Clinical Status	T helper 1–20: 12 T	13 V	21–49: 26 S	27 V	29 D	35 S	38 Y	40 E	T helper 50–69: 51 H	55 L	56 R	57 Q	58 A	59 I	63 G	64 E	66 M	67 T	68 L	70–73: 70 T	73 G	B cell 74–89: 74 V/G	77 E	79 P	80 A	81 S	87 S
1A[1]	N																												
1A[2]	C		S	M						D T	N									F			S					A	
1B	C	Remission	S							D						V	D		I	F			S					A	N
2A	N																												
2B	C	Remission												H	V			D											
3A	N																												
3B	C	Remission																D		I	S						T		
4A	N								A																				
4B	C	Active	S							D															Q	Q	P		
5A	N																	D		I	S						T		
5B	C	Active																D		I	S						T		
6A	N																										T		
6B	C	Active																D					V				T		

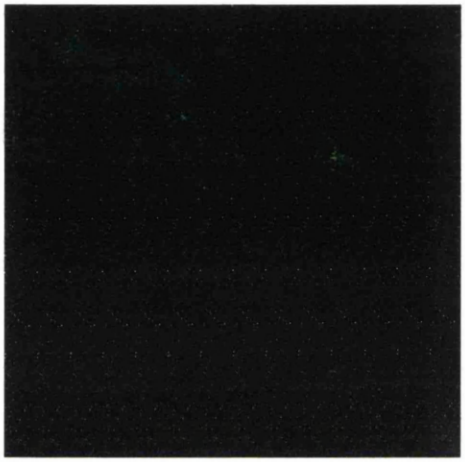
ent e	*HBc Ag* Loc.	Clinical Status	T helper 1–20: 12 T	13 V	21–49: 26 S	27 V	29 D	35 S	38 Y	40 E	49	T helper 50–69: 51 H	55 L	56 R	57 Q	58 A	59 I	63 G	64 E	66 M	67 T	68 L	70–73: 70 T	73 G	B cell 74–89: 74 V/G	77 E	79 P	80 A	81 S	87 S	89
7A	NC		S								T																		Q			
7B	NC	Remission	S					A	E																				Q			
8A	N					T		Q	F	D				I											V							
8B	C	Active				T		L	F	D				I							N				G	Q		I				
9A	N									D															A							
9B	NC	Active	S							D							V		D						A							
10A	C									D							V		D						A							
10B	N	Active		L				A		D					H	V	V		D													
11A	N									D				I																		
11B	C	Active								D				I																V		G
12A	N																															
13A	N																															
14A	N																															

Patient																								
						B cell								B cell						C-terminus				
Code	90				106	107			118	119	127	128			135	136		149	150						
	92	93	96	97			109	113			R		130	131						153	155	156	160	180	181
	N	M	K	F			T	E					P	A	P					G	S	P	T	E	
1A[1]																									
1A[2]																								P	A
1B				I				D																P	A
2A																									
2B																							A		
3A																									
3B																									P
4A																									
4B															P										P
5A																									
5B																								K	
6A													Q							C	T				
Patient																									
						B cell								B cell						C-terminus				
Code	90				106	107			118	119	127	128			135	136		149	150						
	92	93	96	97			109	113			R		130	131						153	155	156	160	180	181
	N	M	K	F			T	E					P	A	P					G	S	P	T	E	

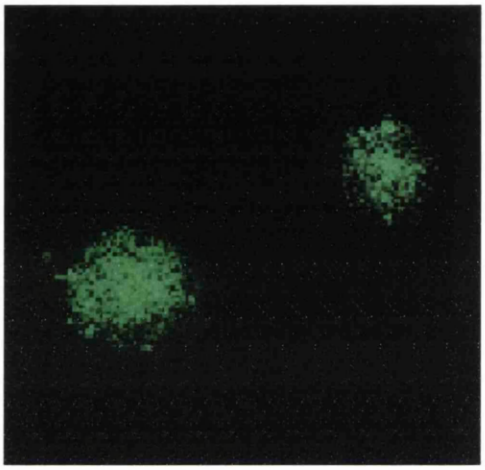
Patient Code	90	92	93	96	97	106	107	109	113	118	119	127	128	130	131	135	136	149	150	153	155	156	160	180	181
							B cell						B cell							C-terminus					
		N	M	K	F			T	E			R		P	A	P				G	S	P	T	E	
6B		V												Q						C	T				
7A								M																	P
7B												L		Q						C	T				
8A																				C					
8B					E											T						L			
9A																									
9B			V							Q				Q											P
10A															P		F								
10B																									
11A																S			R	Q	T				
11B			T													Q		C		Q	T				
12A																									
13A																									
14A																									

Note: B cell and T helper epitopes and C-terminus region areas with their boundaries and wild type variants indicated at top. Amino acids are described by single letter code and numbered from the beginning of HBcAg. Only positions at which changes occurred are shown, so relative proportion of epitopic to non-epitopic areas is skewed in favour of regions where substitutions occurred. Clinical status of 11 later samples from anti HBe positive patients indicated as either in remission or active phase. N, predominantly nuclear; C, predominantly cytoplasmic ; NC, mixed localisation.

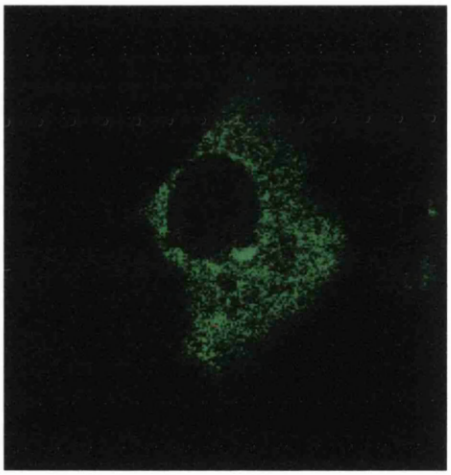
Fig 4.1.2- Intracellular localisation of HBcAg using confocal microscopy. Immunoflourescence staining of HBcAg in Cos-7 cells transfected with constructs containing clones of : A) initial and B) later samples of patients with chronic HBV infection, C) individual HBeAg positive samples as described in the text.



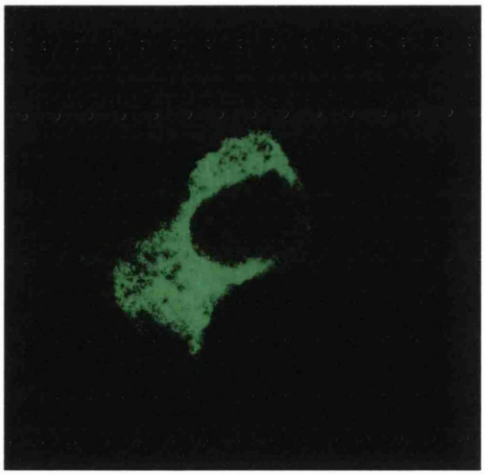
Negative Control



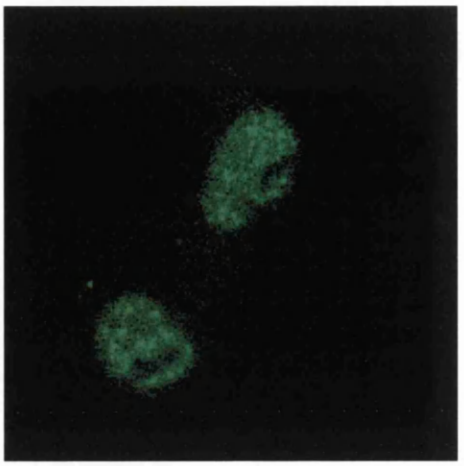
1A[1]



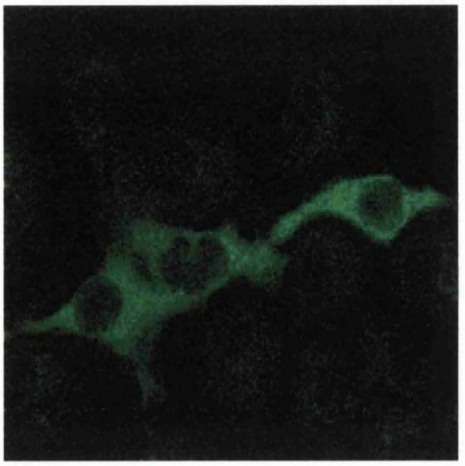
1A[2]



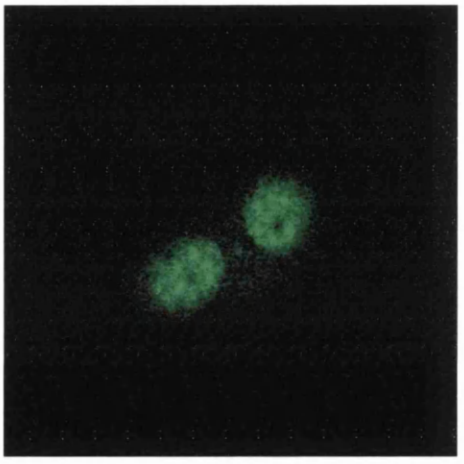
1B



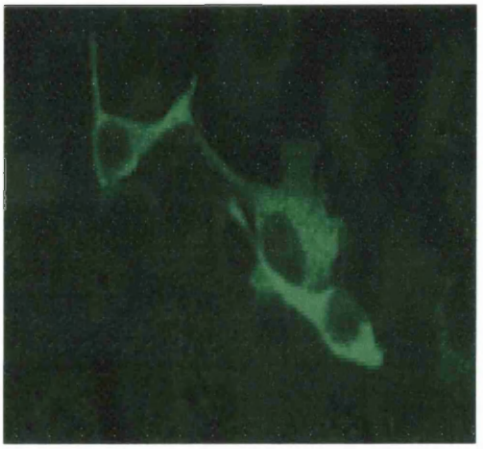
2A



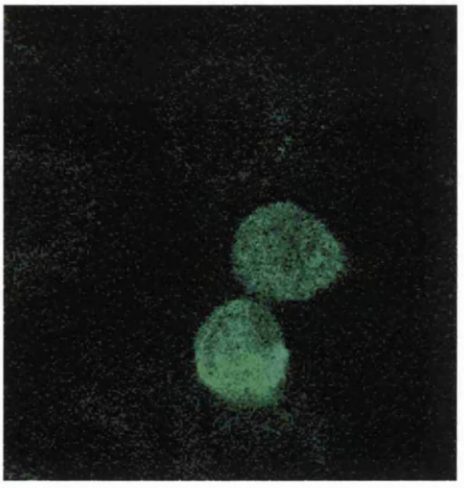
2B



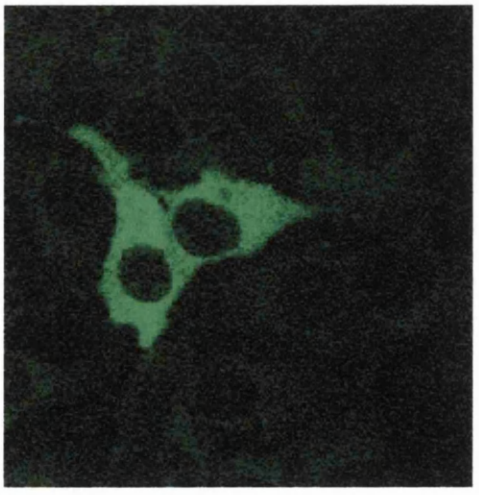
3A



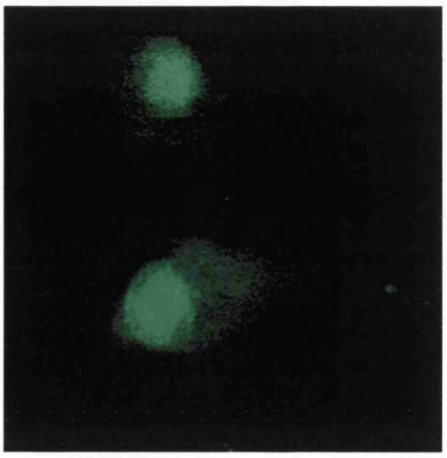
3B



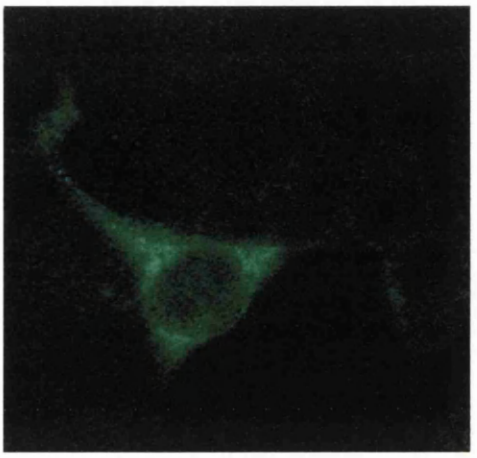
4A



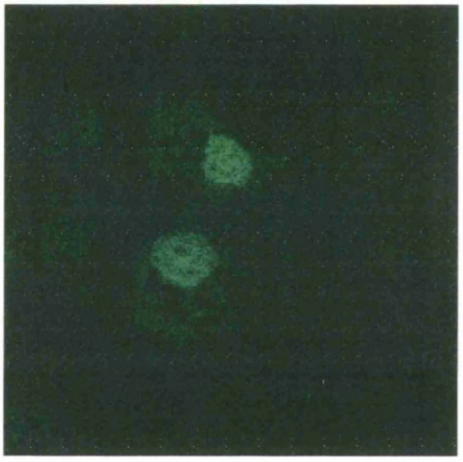
4B



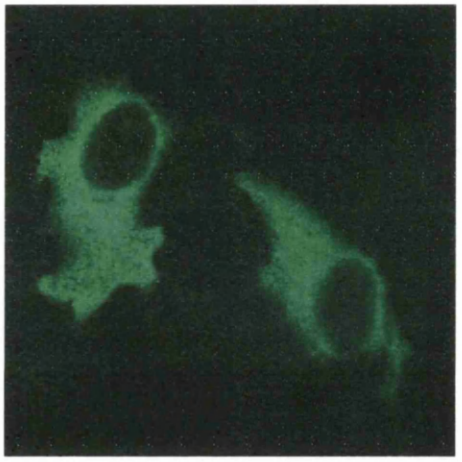
5A



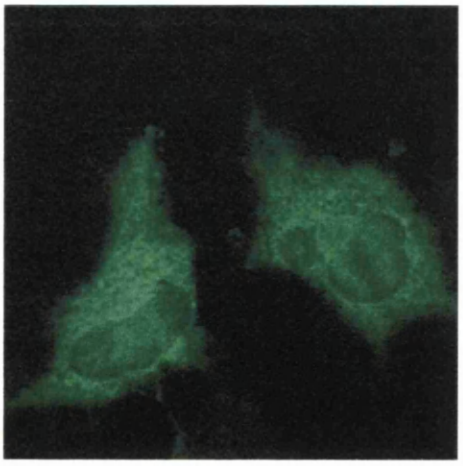
5B



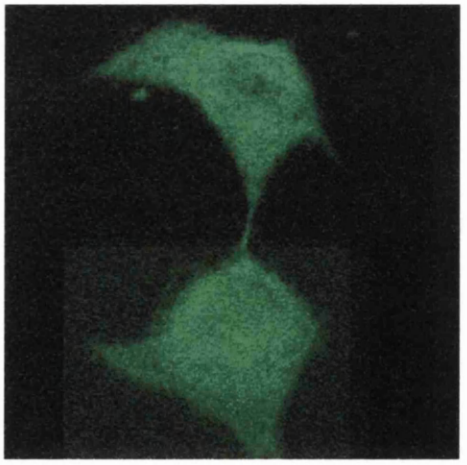
6A



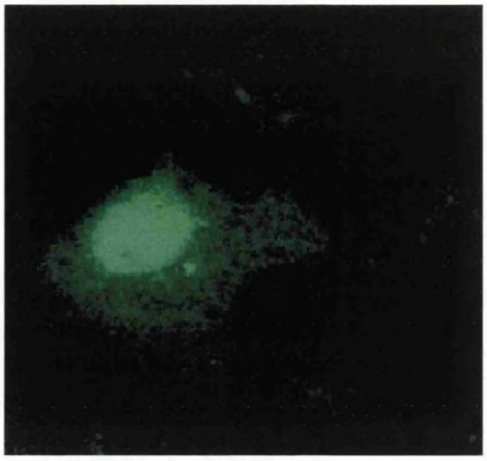
6B



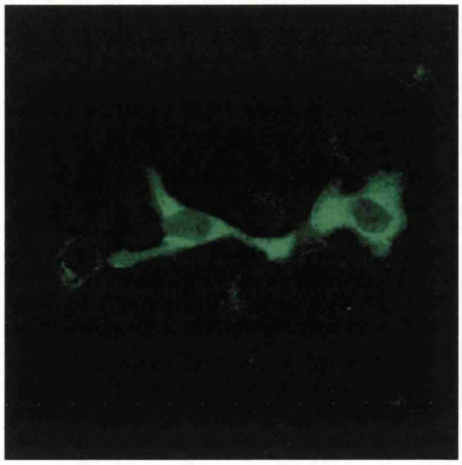
7A



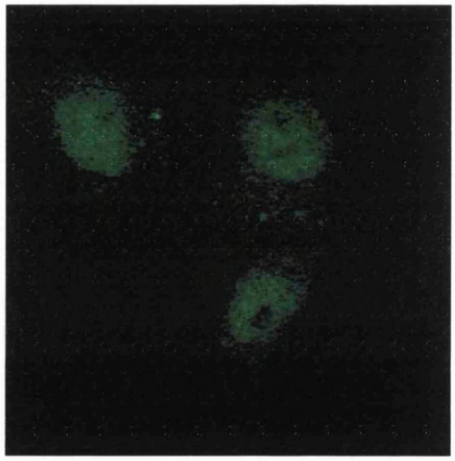
7B



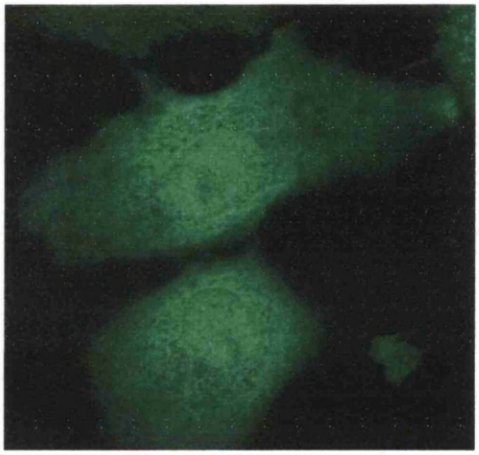
8A



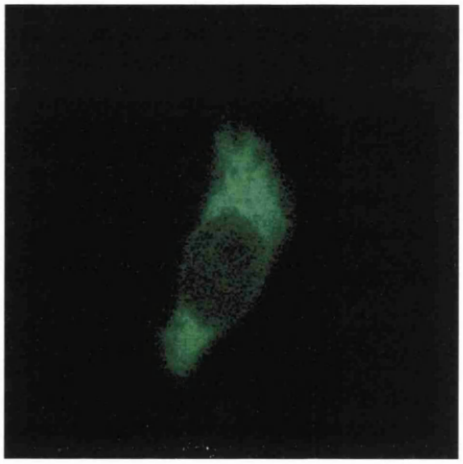
8B



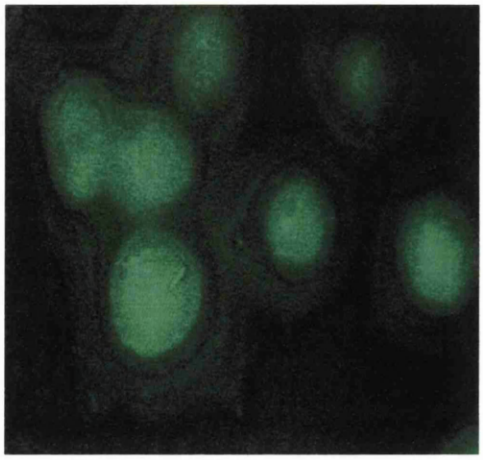
9A



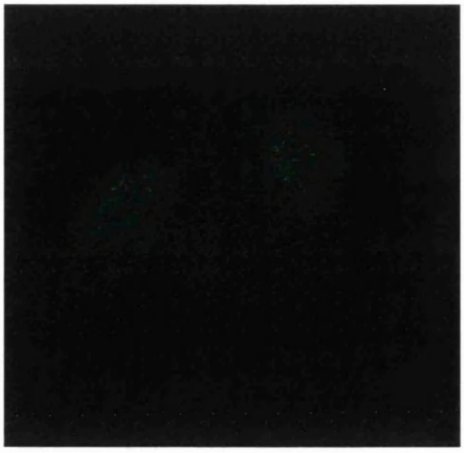
9B



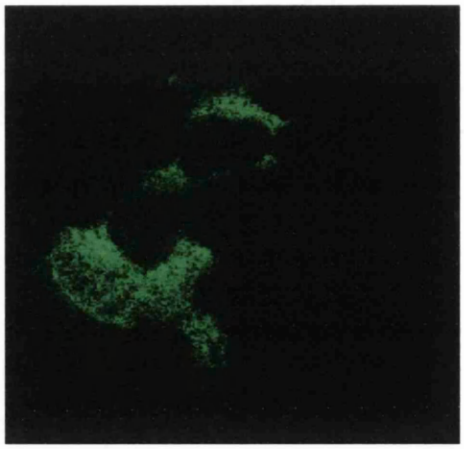
10A



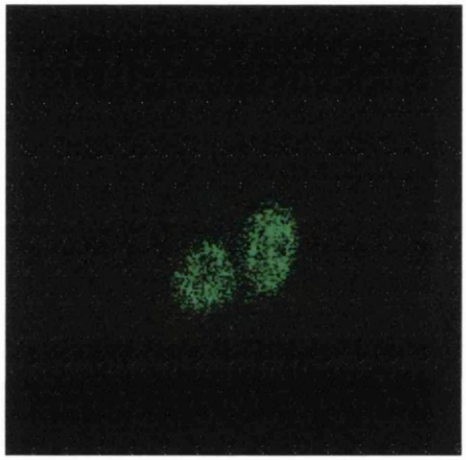
10B



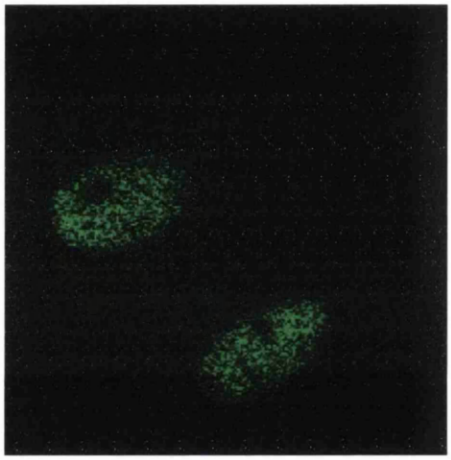
11A



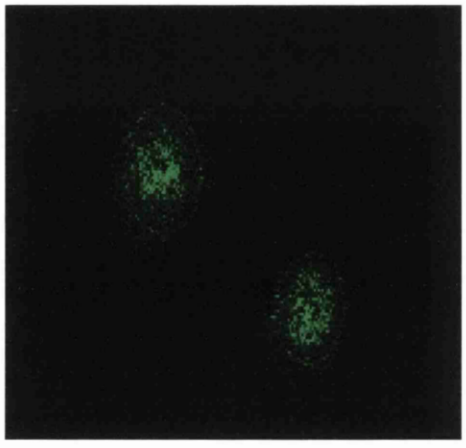
11B



12C



13C



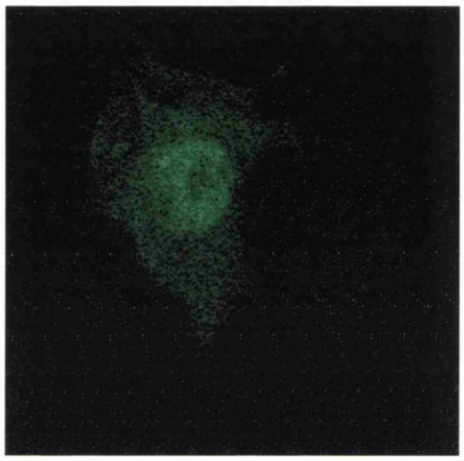
14C

Table 4.1.3 In vitro mutation of cytoplasmically distributed sequences.

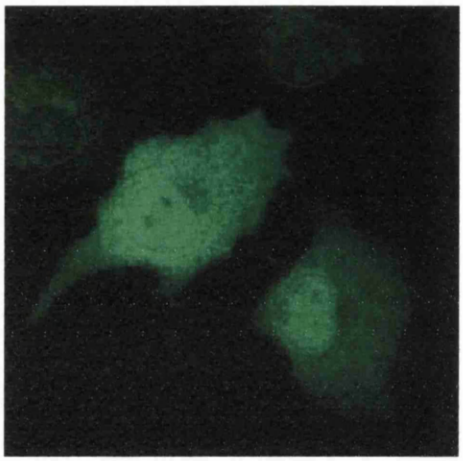
Mutated Virus	Mutagenesis Position B-cell Epitopes HBc-1	HBc-2	C-terrminus	HBcAg Localisation N	N&C	C
2B			T160A	+++	+	
3B	A80T			+++	+	
3B			S181P	++	++	
3B	A80T		S181P	+	+++	
4B	E77Q, P79Q, A80P			+	+++	
4B			S181P	+++	+	
4B	E77Q, P79Q, A80P		S181P	+++	+	
5B			Q177K	+++	+	
5B	A81T		Q177K	+++	+	
10A	V74A			++	++	
10A		A131P		++		++
10A	N74A	A131P		+	+++	
11B		P135Q		+	+++	

The mutations generated are listed according to their amino acid position within an epitope. HBc-1, residues between aa position 74-89; HBc-2 residues between aa position 128-135. +, represented 25% of stained-cells. Letters before each number indicates the mutated sequence to wild type and letters after numbers indicates the original sequence before mutagenesis.

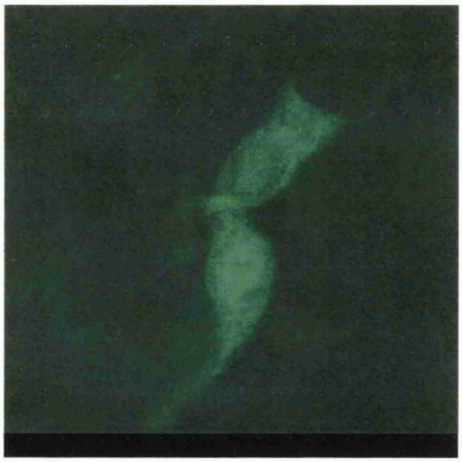
Fig 4.1.3 Localisation patterns of HBcAg in Cos-7 cells transfected with constructs containing mutagenised clones in the B-cell epitopes, C-terminus, and both C-terminus and B-cell epitope regions.



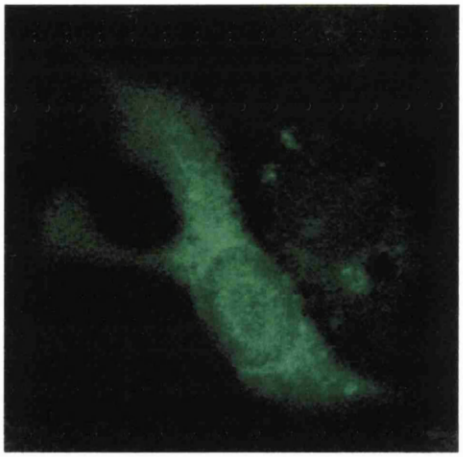
2B. C-terminus mutagenesis, T160A.



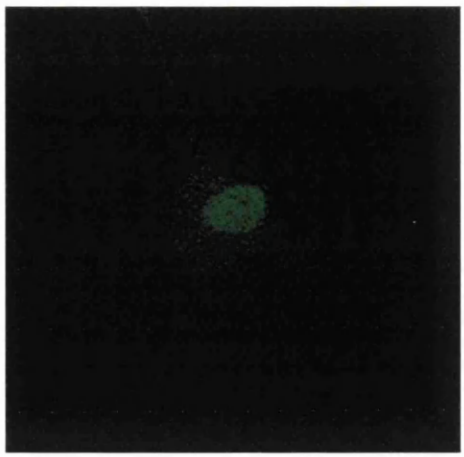
3B. B cell epitope mutagenesis, N74A.



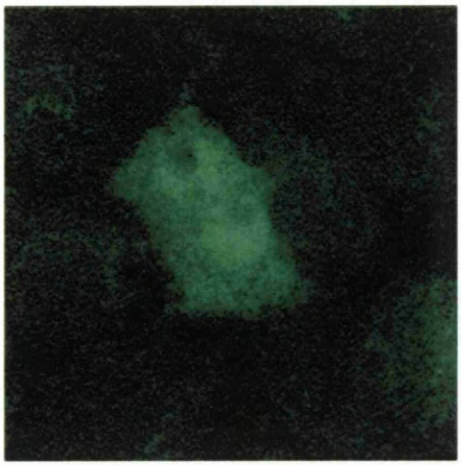
3B. C-terminus mutagenesis, S181P.



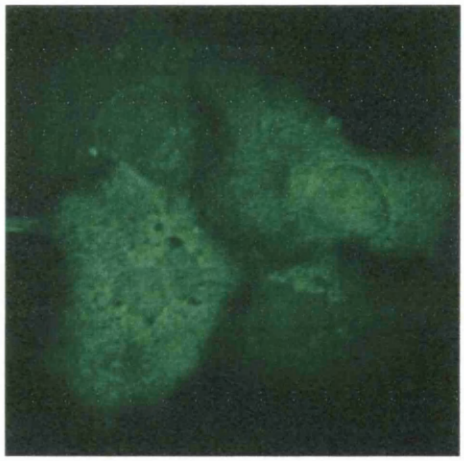
3B. Double mutagenesis, N74A and S181P.



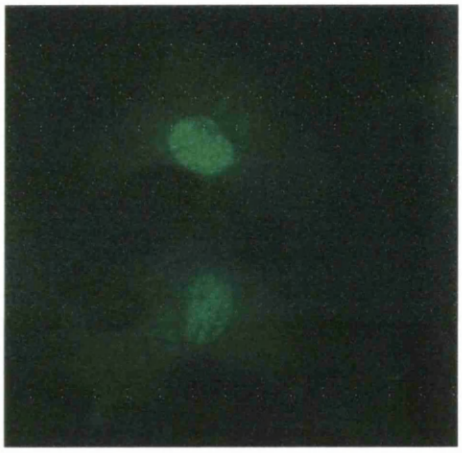
4B. C-terminus mutagenesis, S181P



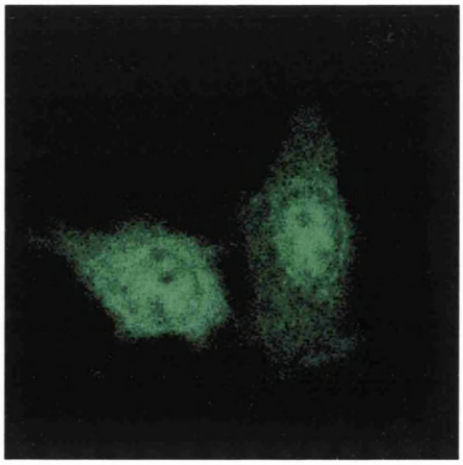
4B. Double mutagenesis, E77D, A79Q, S80P and S181P



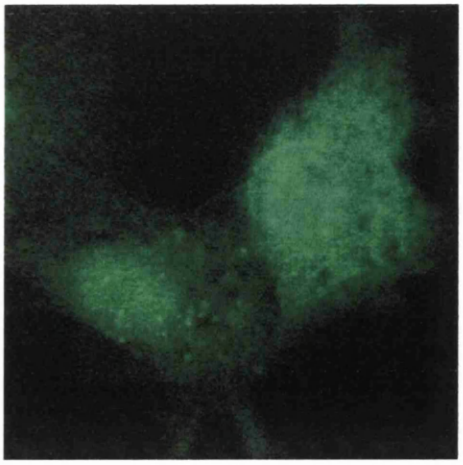
4B. B cell epitope mutagenesis, E77D, A79Q, S80P



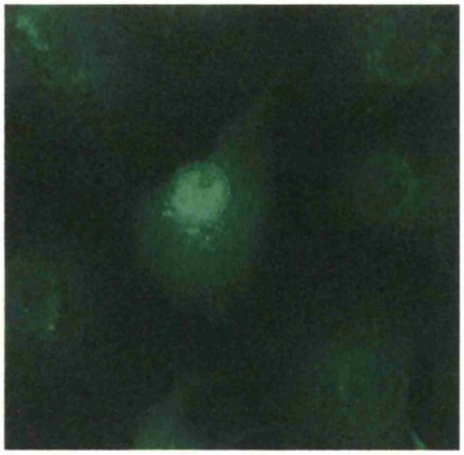
5B. C-terminus mutagenesis, Q177K



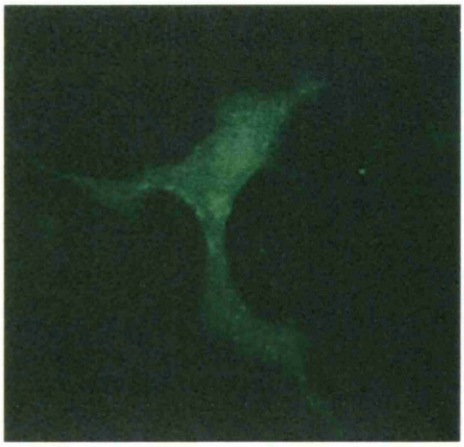
5B. Double mutagenesis, A81T, Q177



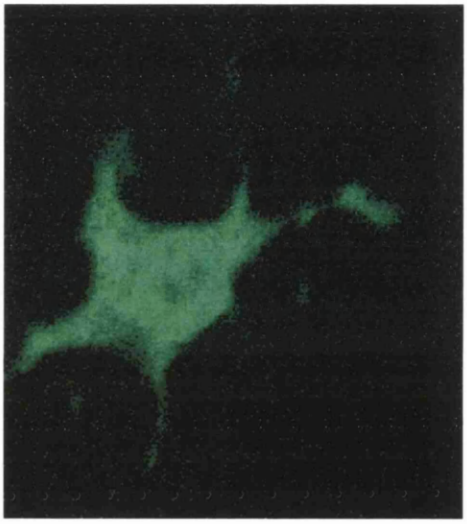
10A. B cell epitope mutagenesis, V74A



10A. Double mutagenesis (both B cell epitopes), V74A and A131P

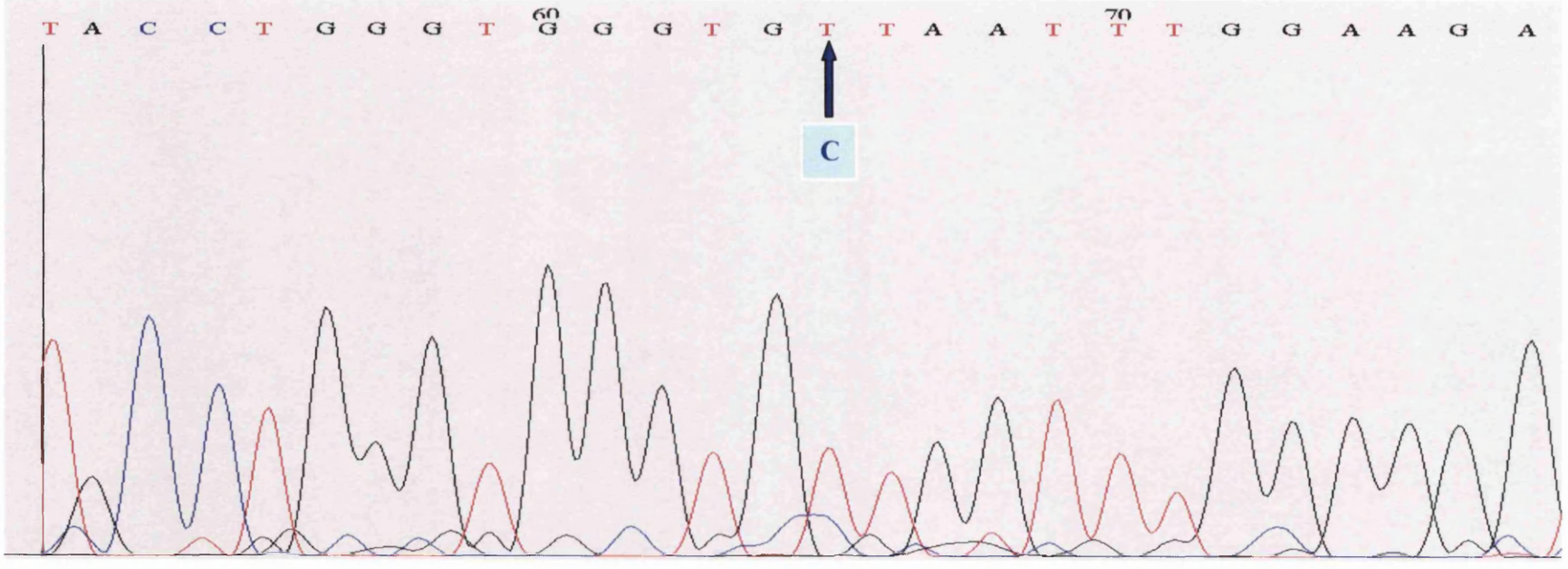


10A.B cell epitope mutagenesis, A131P



11B. B cell epitope mutagenesis, P135Q.

Fig 4.1.4 An Example of HBcAg in vitro mutagenesis.



10A

Note: Electropherogram shows the nucleotide sequence reverted back to original (wild type) sequence (in HB-c1 region V74A) using mutagenesis method. The same changes carried out for double mutagenesis at the same time, e.g., either B cell or C-terminus mutagenised samples used as a template for second mutagenesis.

4.2 Discussion

The entire HBV genome is extremely stable in the early high-replicative HBeAg positive phase, when the immune response is considered to be weak (Hannoun et al., 2000a). However, after HBeAg seroconversion to anti-HBe, multiple clusters of mutations appear in different regions of HBcAg, in particular immunodominant epitopes (see chapter 1.4.4, figure 1.7 and table 1.1). Several studies have shown that amino acid variability of HBcAg is related to the severity of liver disease, both in the HBeAg and anti-HBe phases (Wakita et al., 1991; Ehata et al., 1992; Chaung et al., 1993; Carman et al., 1994, 1995b, 1997b). There are reports describing HBcAg distribution in the cytoplasm, nucleus, or both, in infected or transfected cells (Yamada et al., 1977; Brechot et al., 1980; Gowans et al,. 1983; Uy et al., 1986; Akiba et al., 1987; Hsu et al., 1987; Wu et al., 1987; Nakopoulou et al., 1992; Fang et al., 1994; Bock et al., 1996; Aiba et al., 1997; Naoumov et al., 1997; Chu et al., 1997; Park et al., 1999; Serinoz et al., 2003). Recently, Kawai and colleagues found that hepatocytes exhibiting cytoplasmic HBcAg expression contained more core promoter double mutations (T_{1762} and A_{1764}) than hepatocytes exhibiting nuclear expression (Kawai et al., 2003). Studies on liver biopsies have suggested that the intracellular distribution of HBcAg correlates with disease activity, cytoplasmic expression being associated with more severe disease (Wu et al., 1987; Akiba et al., 1987; Ballare et al., 1989; Tong et al., 1990; Kawai et al., 2003). The cell membrane of HBV-infected hepatocytes may also express HBcAg during chronic infection (Ferrari et al., 1987; Naoumov et al., 1993). However, whether HBcAg intracellular distribution is associated with liver damage is unclear.

Previous studies have been carried out on biopsy material in which the additional influence of host factors makes it difficult to delineate the mechanisms responsible for the intracellular distribution of HBcAg. The present study has confirmed previous observations that there is a close relationship between the topographic distribution of HBcAg in vitro and the inflammatory status of the liver in vivo. The staining pattern is unlikely to be caused by different transfection efficiencies, because the amount of plasmid DNA used was standardized and the experiments in some instances repeated more than two times. For the first time, it has been shown that the distribution of expressed HBcAg is strongly linked to the viral sequence. The presence of cytoplasmic expression of HBcAg correlated with the existence of mutations in the B-cell epitopes and/or carboxy-terminus of the protein. These experiments were carried out in the presence of aphidicolin to block cells in G1/S; consequently, any shift to cytoplasmic expression was not a result of nuclear membrane dissolution. However, to prove that sequence was critical in HBcAg localisation, reversion of C-terminus and/or B-cell epitopes mutations back to the wild type was performed: in most cases, this shifted HBcAg localisation back to the nucleus. Some possible explanations for this shifting follow.

First, only phosphorylated core protein specially targets the nucleus via signal sequences in the C-terminus (Eckhardt et al., 1991; Liao et al., 1995), which interact with kinase or other cellular factors (Hui et al., 1999). Removal of the last 36 C-terminal aa completely abolished phosphorylation of the core without interfering with particle formation (Schlicht et al., 1989). Consequently, altered amino acid residues in close proximity to SPRRR motifs may influence the behaviour of HBcAg in that they may induce a

conformational change, suppressing phosphorylation by ap34^{cdc2} kinase. HBcAg would then remain in the cytoplasm. Although this work was carried out on DHBV core particle, for which size and position of the phosphorylation sites differ from human HBV, the concept that phosphorylation alters the localisation of core protein still remains valid.

Second, Machida and co-workers reported an antigenic site within the C-terminus region called HBicAg (hepatitis B inner cAg, aa residues 148-160) as a potential B cell antigenic determinant (Machida et al., 1989, 1991; Sallberg et al., 1994). Here, 13 out of 22 C-terminus mutations occurred within aa residues 148-160 (Table 4.1.3). If the arg-rich C-terminus can stimulate a B-cell response in humans. As this region is also known to affect nuclear localisation, mutations would explain the shift of HBcag to the cytoplasm. This is in line with observations that escape mutations appeared in the B cell epitopes in patients with ongoing disease (Carman et al., 1997b), as double mutagenesis of both a B-cell epitope and C-terminus residue: all samples shifted back to both nuclear and cytoplasmic and/or predominantly nuclear.

Although B-cell epitope mutations tended to shift HBcAg to the cytoplasm, this was not as effective as C-terminus mutations, which probably direct HBcAg intracellular trafficking by a different mechanism(s). A majority of B cell mutagenised samples localised into the both nucleus and cytoplasm, while all C-terminus mutagenised samples showed predominantly nuclear distribution. This was also observed in 2 of 3 double mutagenised samples. The reason underlying this phenomenon is unclear and deserves more investigations.

We clearly showed, in samples with both nuclear and cytoplasmic localisation, that HBcAg could be associated with the nuclear membrane. This unusual pattern has been

reported before (Yamada et al., 1977; Bock et al., 1996; Wu et al., 1996). As there is no sharp boundary between G1 and S phases, transport of HBcAg to the nuclear membrane may occur during the transition. It may be that nuclear/cytoplasmic localisation may be a transition step between the two compartments (Akiba at al., 1987). Alternatively, there may have been high levels of nuclear HBV DNA with inefficient transportation from the nucleus. However, there is no strict division between predominantly nuclear and predominantly cytoplasmic core protein localisations in vivo, because core particles are usually present in both compartments (Yamada et al., 1977; Naoumov et al., 1997).

In a majority of cases, a predominantly cytoplasmic distribution in vitro was found in both samples from patients in remission and with active disease. These samples ($1A^1$, $1A^2$, 2B and 3B) selected more mutations in Th epitopes than in C-terminus region and/or B cell epitopes (Table 4.1.2). Of the seven patients with active disease six showed a shift from nuclear to cytoplasmic expression between the earlier and later sample of which, four had a higher proportion of mutations in B cell epitopes than other regions (Table 4.1.2, 4B, 9A, 10B and 11A). All samples with active disease were anti-HBe positive, of which 4 showed predominantly nuclear staining (see above). In keeping, in 4 samples from continuously anti-HBe positive patients with predominantly nuclear staining, all were taken from patients with active disease. The only apparent correlation in HBcAg sequences between this group (samples 8A, 9A, 10B and 11A) and other anti-HBe positive with predominantly cytoplasmic staining was that fewer mutations in these samples have been found compared to later ones with predominantly cytoplasmic distribution (Table 4.1.2). Although this unusual expression has been reported before (Wu et al., 1996; Borg et al., 2000; Kawai et al., 2003), it is likely that after HBeAg

seroconversion to anti-HBe there is co-existence of both antigen and antibody for some time (Maruyama et al., 1998), with concurrent persistence of low concentration of integrated HBV DNA in the liver after the disappearance of HBeAg (Wu et al., 1996). Nevertheless, various studies reported continued liver disease activity after conversion to anti-HBe status in certain patients that might be related to continued viral replication and secretion with the presence of HBcAg in the nuclei. Even in some patients with cirrhosis and severe chronic hepatitis, HBcAg has been reported to be predominantly nuclear (see chapter 1.4.1F). The reason underlying this pattern is not clear, but might be related to the specific patient population studied (Hadziyannis et al., 1983).

We believe that the differential subcellular distribution of HBcAg in chronic HBV infection in different reports might be related to genotype-specific sequence variability between different ethnic groups indeed. This is currently under investigation but would be in keeping with recent data on different clinical outcomes in some East Asian patients infected with either genotype B or C (Lindh et al., 1999; Kao et al., 2000a and b, 2001; Orito et al., 2001; Yoo et al., 2003).

In summary, we have shown that subcellular expression of HBcAg is related to serology and depends on HBcAg sequence even in sequential samples from individual patients. C-terminus and/or B cell epitope mutations are involved in the shifting of HBcAg between nucleus and cytoplasm. These results were generated using an artificial system. As there is no susceptible cell line that can be truly infected by HBV, the biological importance of such variants on the viral life cycle and cell behaviour are still unclear.

If expression of HBcAg is predominantly cytoplasmic, it is conceivable that at least some of these processed peptides would express on the cell membrane (Saito et al., 1992) (see

chapter 1.4.3F), which renders this target viral antigen susceptible to immune attack and severe liver injury. Conversely, if HBcAg is localised to the nucleus, then little or no HBcAg should be expressed on the cell surface with a concomitant decrease in liver cell injury.

CHAPTER 5 HBV In Vitro Infection

5.1 Aims of study

Because of the unavailability of a relatively susceptible cell line that can be infected by hepadnaviruses, there is no readily available system for studying the early events of the HBV life cycle. Thus, most studies are carried out by transfection. FTO 9.1 cell line (a rat hepatoma cell line which is transfected with a construct containing human-Annexin V, h-A V) was employed for HBV infection. The effect of varying concentrations of DMSO and recombinant h-A V was assessed. Cells and culture supernatants were assayed at various times post infection by both IF microscopy (HBcAg staining in nucleus) and PCR of culture supernatants. Supernatants from these initially infected cells were then used to infect fresh FTO 9.1 cells with similar outcomes to primary infection. The virus has been further propagated in this manner in at least one further passage and can be reliably reproduced thus indicating the value of this cell line for in vitro studies of HBV.

5.2 Introduction

HBV displays a distinct hepatotropism and a narrow host range in vivo (Ganem et al., 1987, 1994). However, very little is known about the interaction of HBV with its host cells, mainly because of difficulties in the development of suitable tissue culture system. In vitro studies of HBV infection system involve chromosomal replication of HBV DNA which does not include the early events of infection. Several studies using transfection and/or transgenic technology have suggested that the species barrier of HBV infection and replication may be located at the early step of viral adsorption (Gong et al., 1999; Paran et al., 2001). Transformed cell lines would be of interest for in vitro culture of HBV, however, they are also thought to lose some of their liver specific markers and functions, so losing susceptibility to HBV. The potential development of an in vitro culture system for HBV replication provided an opportunity to investigate molecular and cellular events in cloned

cells (Roingeard et al., 1990b). However, attempts to establish a stable and reproducible in vitro system for production of HBV by infection of cells have not been successful.

5.2.1 Primary Human Hepatocyte Cells

Primary human hepatocytes are susceptible to viral infection as shown by de novo production of viral proteins, appearance of HBV DNA and DNA polymerase activity, and HBV-specific transcripts. However, with increasing time in tissue culture, such cells gradually lose the capability for production of viral proteins. Also, there is loss of the "hepatocyte-like " morphology, presumably the result of a loss of differentiated functions of the hepatocytes in cell culture (Galle et al., 1994). A further problem is that primary human hepatocytes are very difficult to obtain. By adding chemicals (such as DMSO, PEG), infectivity can be stabilised and life and viability of the cells can be extended. However, as the addition of these artificial substances does not mimic the in vitro process and may be affecting the infection mechanism, it is desirable to find another system.

5.2.2 Human Liver Tumour Cell Lines

Hepatoma cell lines can be characterised by their expression of hepatocyte- and biliary-specific genes (Zvibel et al., 1998). It is possible to study HBV production in culture by transient transfection of human hepatoma cell lines (Shih et al., 1989): HepG2 (Roingeard et al., 1990b), Huh 6-c15 (Tsurimoto et al., 1987), Huh-7 (Yaginuma et al., 1987) and Hepa RG (Gripon et al., 2002).

However these cell lines do not support stable HBV infection. It is postulated that human hepatoma cell lines do not have the receptor for virus attachment, penetration, and fusion leading to productive infection (Bchini et al., 1990). On the other hand, some infected HepG2 cell line have been achieved and reported to release HBV particle into the medium with the same antigenic, morphologic, and biologic properties as HBV virions (Bchini et al., 1990; Paran et al., 2001).

5.2.3 FTO Cell line and Annexin-V

Chapter 1.2A details the existence of a receptor-ligand relationship between hA-V and HBsAg. In this regard, transfection of hA-V into a non-susceptible rat hepatoma cell line, FTO2B, resulted in conferred susceptibility to HBV infection. FTO 9.1 (Gong et al., 1999). It is interesting to note that rat Annexin V does not bind to HBsAg, although it has a sequence homology of more than 90% as compared with hA-V at the amino acid level (Gong et al., 1999). This cell line can therefore be considered both susceptible and permissive to HBV replication, although FTO-hA-V cells may lose expression of Annexin-V over time (Dornan et al., unpublished results).

5.3 Objectives:

1-Which diagnostic procedure can be used to confirm true FTO 9.1 infection i.e., episomal, and not chromosomal, replication.

cccDNA specific PCR of cell extracts is a marker of internalization of virus by the cells, and confirms infection. HBcAg staining would also indicate expression of HBcAg by the cells.

2-How many viral passages can be achieved in this cell culture system?

3-What are the effects of chemical additives on the infectivity of FTO 9.1?

5.4 Methods

5.4.1 Cell Culture and Infection

FTO 9.1 cells (Innogenetics, Gent, Belgium) were grown at 37 C, 5% CO_2 in Hams F12: Dulbeccos modified Eagles medium (1:1 v/v) containing 10% foetal bovine serum, 100 IU/ml Penicillin, 100ug/ml Streptomycin and 2mM L-Glutamine (Life Technologies, Paisley, UK), and 600ug/ml G418 (The media for negative control did not contain G418). Cells were passaged every 3-4 days. Cos-7 cells were grown in DMEM with the same supplements as for the FTO 9.1 cells, except that G418 was not added to the medium. In our infection experiments the FTO-2B cell line (rat hepatoma without annexin V) was not

available for study and for control cell lines, we used COS-7 cells which should not be permissive.

For infection, cells were plated at $3x10^5$ cells in 24 well plates. Infection was carried out for 16 hrs with HBeAg positive serum containing ~$5x10^6$ genome equivalents (geq)/ml. HBV positive serum was diluted in FCS and media to give multiplicity of infection (MOI) of 100. At the time of infection either 2% DMSO or rAV (10ug/ml) added to the media. Cells were infected overnight at 37°C, then washed thoroughly (7 times) with PBS, and 200ul of the last wash was extracted (Figure 5.1, a), thus it is unlikely that the positivity of FTO cells seen at day 5 is a result of amplification of DNA from virions present in the original inoculum. For passage of the virus, supernatant of infected cells from day 5 was carefully removed to prevent contamination with cells and the supernatant added to fresh FCS and then added to fresh cells. Cells were examined for fluorescence under a Nikon Microphot-SA microscope. This experiment has been repeated twice with consistent results.

5.5 Results

5.5.1 Cell passage number and infection

Previously, our colleagues infected both early and late passage numbers of the FTO 9.1 cell line, and they found that the earlier passage number, 8 was more infectable than a later passage number, 20 (Table 5.1) (Dornan, unpublished data). As a result of this, we used the cells only at passage number 18 or less.

5.5.2 HBV core gene PCR of cell culture supernatants from FTO9.1 infected cells

Figure 5.1 shows the results of PCR on cell culture supernatants from cells infected with an HBeAg positive serum. A C-gene PCR was employed on the supernatant of infected cells from days 2 and 5 to confirm the infection (Figure 5.1, b, c). 200 ul of supernatant was extracted at day 5 and the supernatant became core PCR positive (Figure 5.1, d).

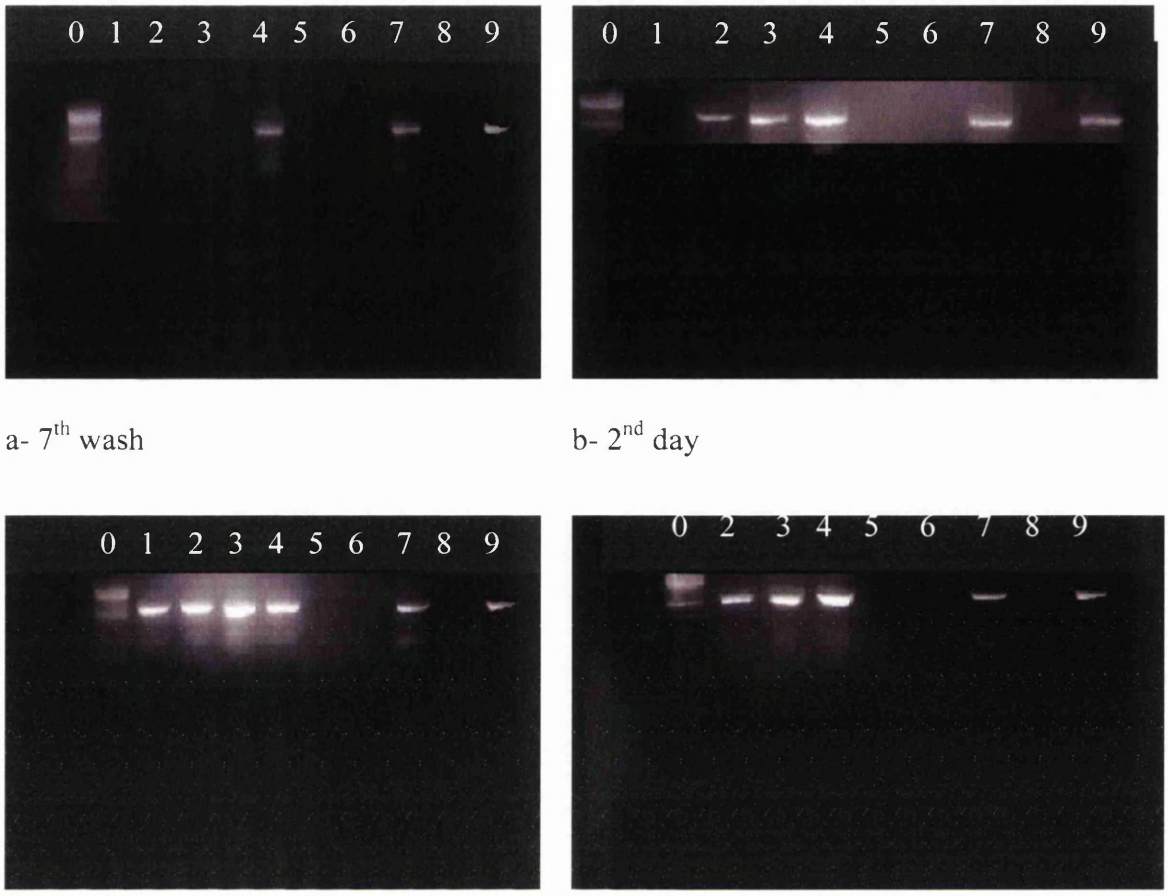
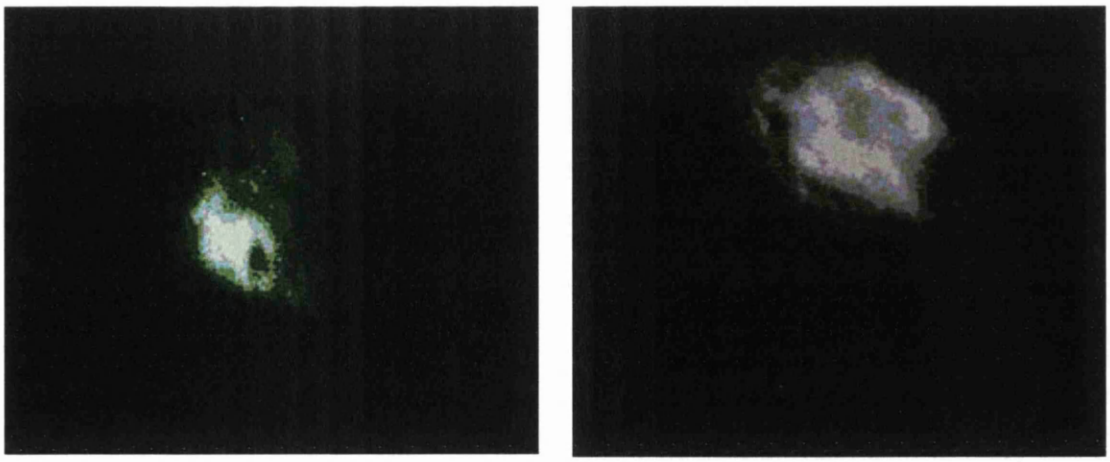


Fig 5.1 Comparison of PCR results obtained from different time intervals during FTO 9.1 cell line infected with HBsAg positive sera. Lane 0, DNA molecular weight marker; lane 1, HBV-contained media without DMSO/hAV; lane 2, HBV positive media plus rAV; lane 3, cells contained DMSO; lanes 4 and 5, DNA extraction positive and negative controls; lanes 6 to 9, 1st and 2nd PCR negative and positive controls, respectively.

The COS-7 cell line remained negative (results not shown). cccDNA-specific PCR had been carried out before (results not shown), using specific primers on each side of the gap, to confirm internalisation of the virus into the cells.

5.5.3 HBcAg staining

Cells grown on 13 mm coverslips in 24 well plates were analysed for the expression of HB core antigen 5 days after infection. HBcAg staining can be clearly seen in the nuclei of the cells (Figure 5.2, A).



A B

Fig 5.2 FTO 9.1 cell line expressing HBcAg using immunofluorescene microscopy.

A) HBcAg staining of 5th day post infection with x40 magnification. B) HBcAg staining of transfer experiment, 5 days after supernatant transferring of previous HBV-infected FTOcells to the fresh cells with magnification of x75. (The poor quality of pictures is due to unavailability of better images from the old files and early training on IF microscopy at the beginning of research course).

5.5.4 Effect of recombinant annexin V and DMSO on infectivity of FTO 9.1 cells

It was noted that the addition of 10ug/ml recombinant annexin V to the cell culture medium resulted in more infection efficiency (Figure 5.1- lane 2; Table 5.2). As DMSO had been shown previously to improve infection efficiency, this was also added to the medium (Figure 5.1, lane 3). In the absence of rAV one of four wells proved positive by PCR at day 2 whereas if 10 ug/ml rAV is present, three of four wells were positive (results not shown).

5.5.5 Transfer Experiment

Supernatant from the fifth day after priming infection was then used to inoculate fresh FTO 9.1 cells. Cells stained positive for HBcAg (Figure 5.2 B). This experiment was carried out at least twice by our colleagues, but Cat III became unavailable, we could not analyse further virus passages.

Table 5.1: Effect of FTO-hAV cell passage number on infection in the presence of rAV and DMSO.

Cell Passage number	Day 2 S gene PCR	Day 5 S gene PCR
Passage 8	Negative	Negative
Passage 8	Positive	Positive
Passage 8	Positive	Positive
Passage 20	Negative	Negative
Passage 20	Positive	Negative
Passage 20	Negative	Negative

Table 5.2- Effect of rAV and DMSO on infection of FTO cells

Additions	Day 2 HBcAg PCR	Day 5 HBcAg PCR
-	Negative	Negative
HBV	Negative	Positive
HBV + rAV	Positive	Positive
HBV + DMSO	Positive	Positive
HBV +rAV + DMSO	Positive	Positive

5.6 Discussion

Several cell lines retaining some of the hepatocyte markers are available that support HBV transcription and replication upon plasmid transfection. However, even these permissive cell lines are not susceptible to HBV infection. An inevitable conclusion is that the early steps of virus-cell attachment and entry determine the viral tropism and susceptibility. These results demonstrate the capacity of the rat hepatoma (FTO) cell line, transfected with human annexin V (FTO 9.1), to be fully infected by HBV. This cell line can therefore be considered both susceptible to and permissive for HBV. Using DMSO increased infectability, but its action is non-specific. It may be possible to increase infection efficiency with manipulation of the growth conditions.

Human annexin V presents itself as a likely candidate for the species specificity determining receptor for HBV, because of the ability of a non-susceptible cell line to become susceptible

following transfection with DNA for this protein (Gong et al., 1999). The effect of rAV on infection may be explained by its competence for binding phospholipids, as HBsAg is extensively lipidated (Neurath et al., 1994). However, binding, and subsequent penetration of the hepatocyte membrane by HBV via HBsAg interacting with annexin V may not explain the hepatotropism of HBV. Clearly, since HBV virions can be produced after transfection of a rat cell line and also in transgenic mice (Farza et al., 1988; Shih et al., 1989; Guidotti et al., 1995), the species barrier for HBV is not at the level of transcription/translation of HBV specific genes. Rat/mice nuclear or cytoplasmic factors may act on HBV promoters/enhancers to achieve replication, thus the species barrier is present at the initial stage of entry of the virus into the host cell. Since h A-V is expressed in most tissues (reviewed by Benz and Hofmann, 1997) this receptor does not account for the hepatotropism of HBV. It is possible that the tissue specificity results from downstream events and only liver specific factors will result in new virus after acting on the infecting DNA. These results should now allow the study of various aspects of the viral life cycle that could only be studied with the use of, e.g. primary hepatocytes, which are difficult to obtain and have limited growth characteristics, in term of retaining their hepatocyte specific functions. In future, further neutralisation assays should be carried out, using various monoclonal and polyclonal antibodies to assess their ability to block HBV infection.

In conclusion, this cell line does appear to be useful in the study of the early stages of HBV infection, and with continued investigation, should be available for the research for more information about HBV receptors and the development of new protective or preventive treatments.

General Conclusion

In conclusion, the data presented in this thesis focused on several important issues. First, C gene variation relates to ethnic background; there are various genotypes associated with specific subtypes, which may have migrated with their human hosts. Peoples with related ethnicity, despite wide geographic location, have a set of common variants. Small changes (one or two substitutions) are selected, either negatively or positively, once there is little intermixing between these people (e.g. once they became a tribe). As more sequence data accumulate from different geographic origins, it is expected that HBV genomic classification will become refined, contributing to the finer mapping of the relation between diverse geographic origins and the distribution of HBV strains. Further, sequence analysis of core gene, particularly, in the CTL restricted regions, provides a useful tool for better understanding of biology of HBV strains, and finally, for design of immunomodulatory therapies upon vaccination.

Second, the dependence of the pattern of HBcAg localisation in vitro (in this thesis) on the sequence of variants and its correlation with the serology of chronic HBV infection, supports the hypothesis that such variants may have biological significance by influencing such things as antigenic expression. The effect of those variants on viral life cycle could be studied by applying a suitable cell culture system, such as the FTO9.1 cell line. As it appears from this work, the data presented in chapters 4 and 5 are somehow speculative. Therefore, applying a suitable cell culture system for HBV infection would be useful in the study of biological behaviour of HBV variants in terms of new protective and preventive treatment.

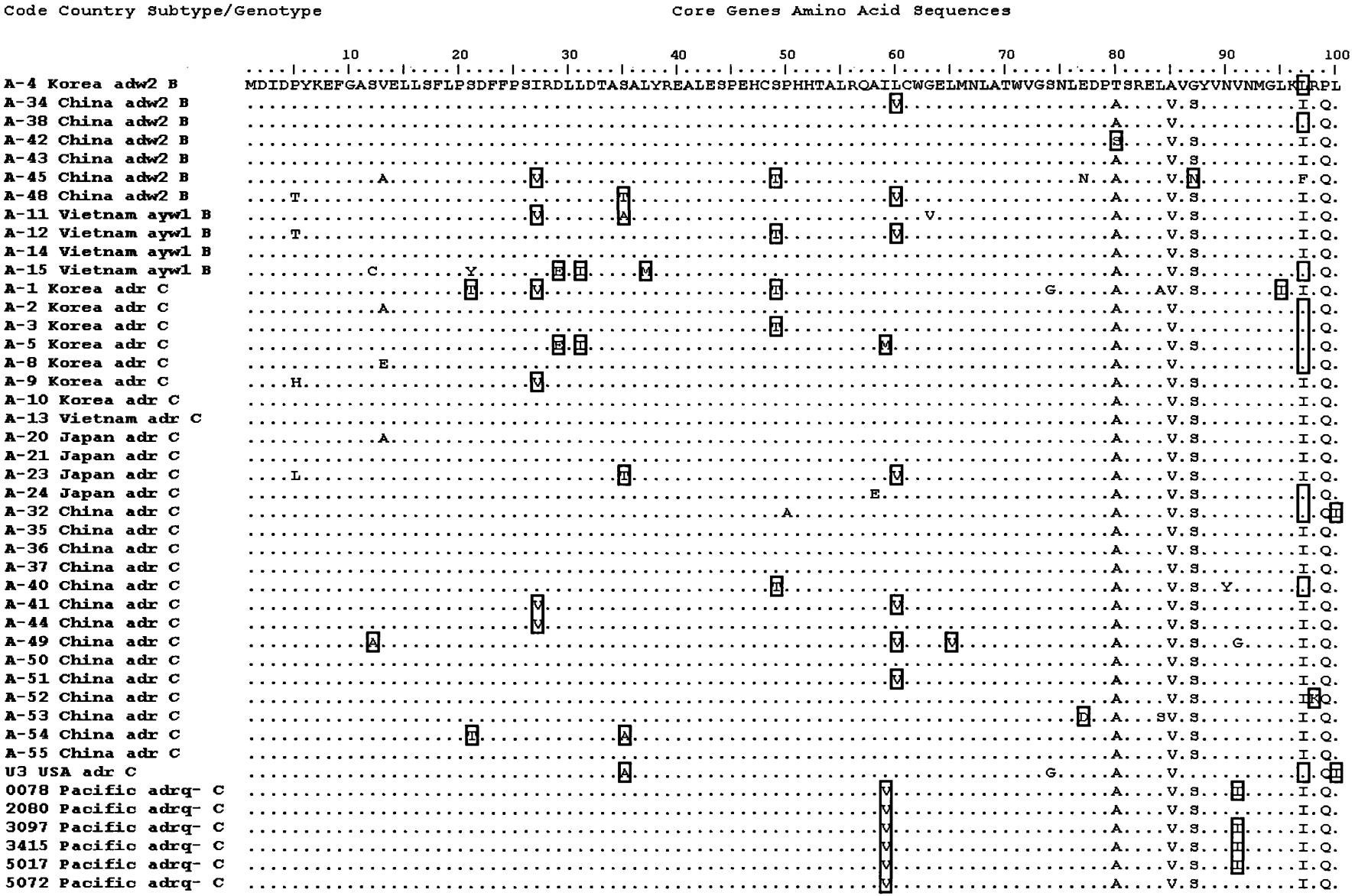
Appendix

A.I

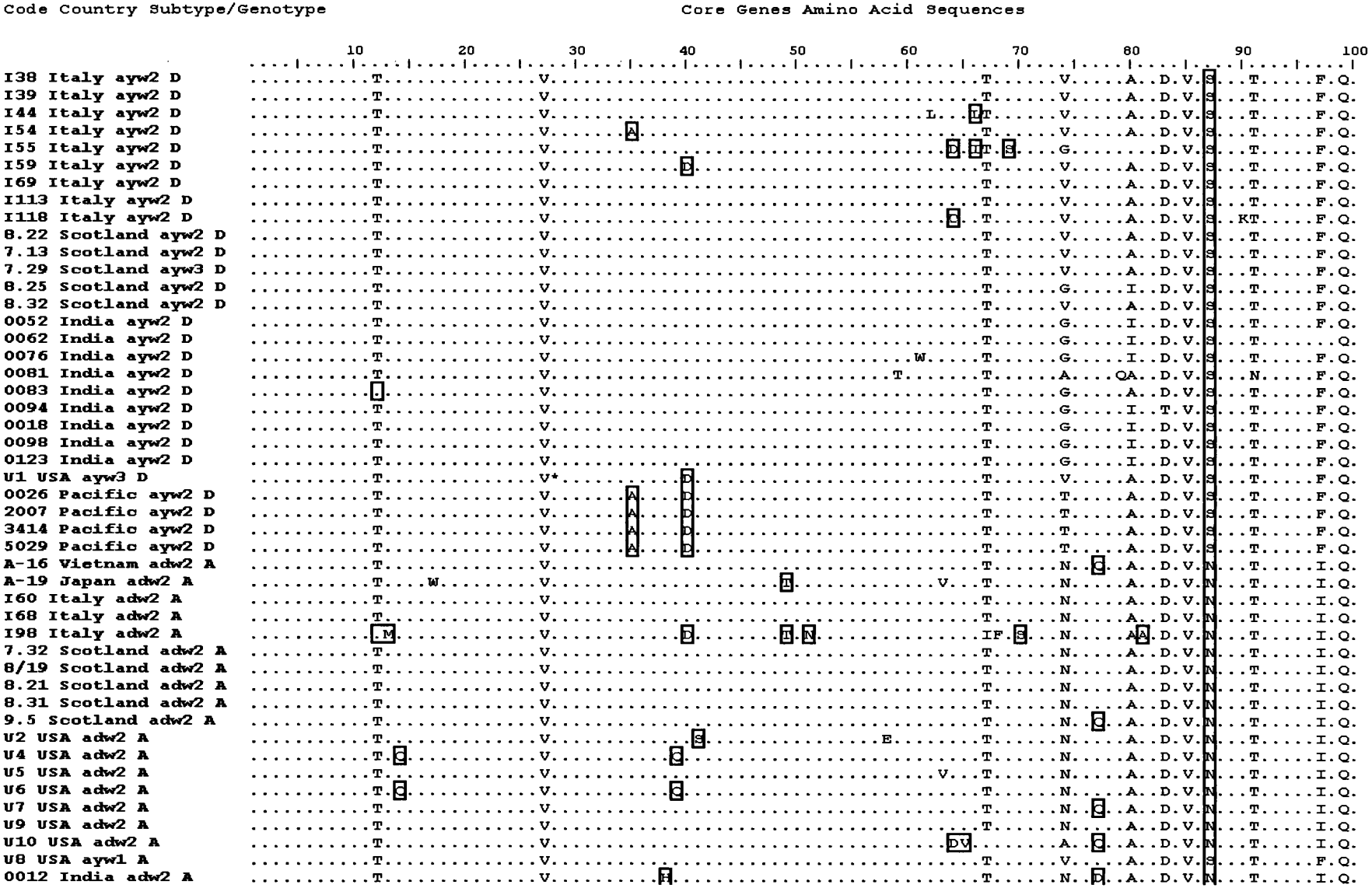
Table A.I.1- Alignment of complete amino acid sequences of HBcAg which shows genotype/subtype identification and geographic origins of 91 sera.

Note: Amino acid residues are numbered from the beginning of the HBcAg using the single letter code. Apart from genotype A, which had 185 amino acids, all the other genotypes had 183 amino acids, therefore using Bioedit software (Clustal W alignment) 2 extra amino acids were added at positions 153 and 154 for a proper alignment.

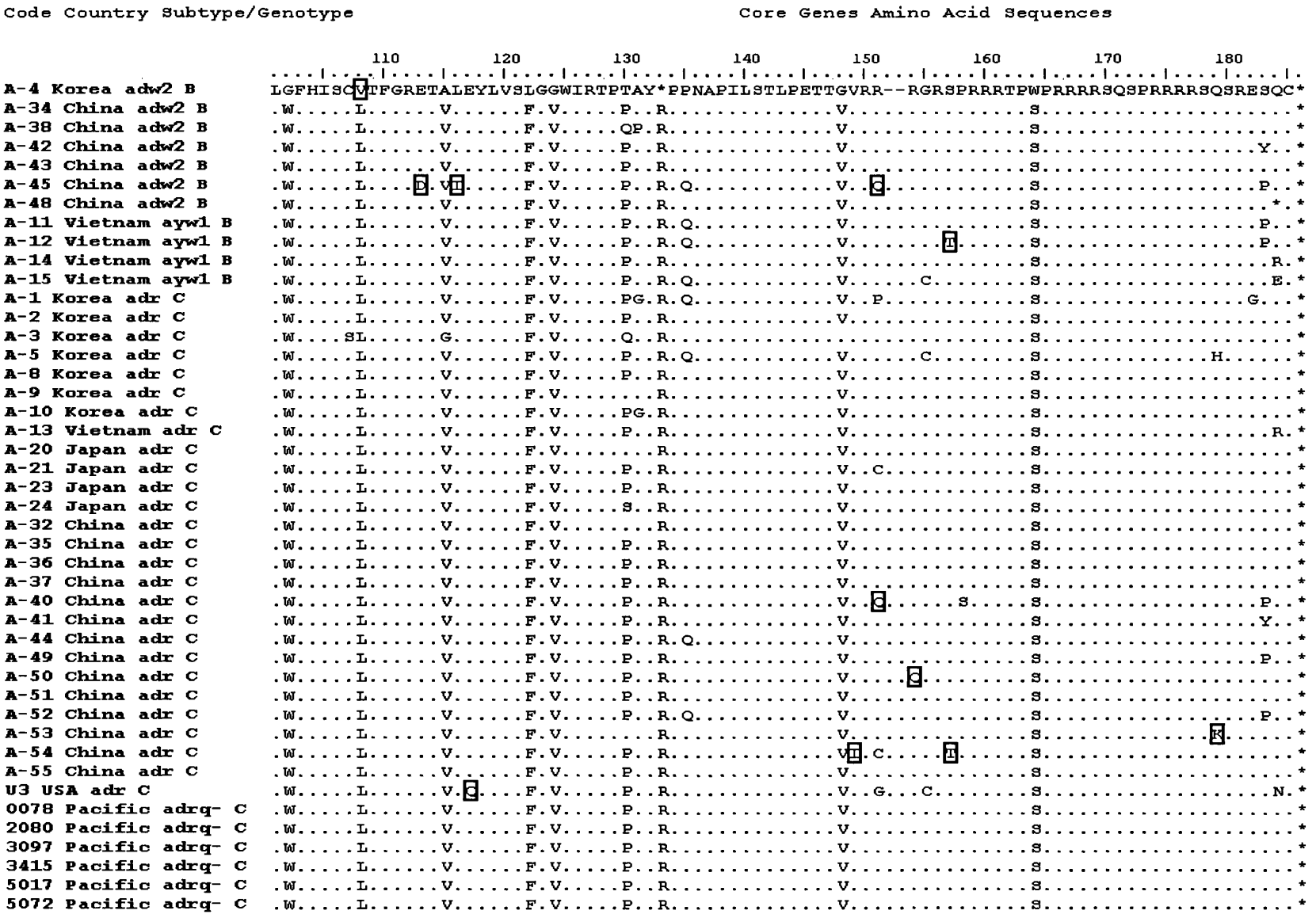
Appendix


Code Country Subtype/Genotype
Core Genes Amino Acid Sequences
10 20 30 40 50 60 70 80 90 100
A-4 Korea adw2 B MDIDPYKEFGASVELLSFLPSDFFPSIRDLLDTASALYREALESPEHCSPHHTALRQAILCWGELMNLATWVGSNLEDPTSRELAVGYVNVNMGLKIRPL
A-34 China adw2 B
A-38 China adw2 B
A-42 China adw2 B
A-43 China adw2 B
A-45 China adw2 B
A-48 China adw2 B
A-11 Vietnam ayw1 B
A-12 Vietnam ayw1 B
A-14 Vietnam ayw1 B
A-15 Vietnam ayw1 B
A-1 Korea adr C
A-2 Korea adr C
A-3 Korea adr C
A-5 Korea adr C
A-8 Korea adr C
A-9 Korea adr C
A-10 Korea adr C
A-13 Vietnam adr C
A-20 Japan adr C
A-21 Japan adr C
A-23 Japan adr C
A-24 Japan adr C
A-32 China adr C
A-35 China adr C
A-36 China adr C
A-37 China adr C
A-40 China adr C
A-41 China adr C
A-44 China adr C
A-49 China adr C
A-50 China adr C
A-51 China adr C
A-52 China adr C
A-53 China adr C
A-54 China adr C
A-55 China adr C
U3 USA adr C
0078 Pacific adrq- C
2080 Pacific adrq- C
3097 Pacific adrq- C
3415 Pacific adrq- C
5017 Pacific adrq- C
5072 Pacific adrq- C

Appendix


Code Country Subtype/Genotype
Core Genes Amino Acid Sequences
10
20
30
40
50
60
70
80
90
100
I38 Italy ayw2 D
I39 Italy ayw2 D
I44 Italy ayw2 D
I54 Italy ayw2 D
I55 Italy ayw2 D
I59 Italy ayw2 D
I69 Italy ayw2 D
I113 Italy ayw2 D
I118 Italy ayw2 D
B.22 Scotland ayw2 D
7.13 Scotland ayw2 D
7.29 Scotland ayw3 D
B.25 Scotland ayw2 D
B.32 Scotland ayw2 D
0052 India ayw2 D
0062 India ayw2 D
0076 India ayw2 D
0081 India ayw2 D
0083 India ayw2 D
0094 India ayw2 D
0018 India ayw2 D
0098 India ayw2 D
0123 India ayw2 D
U1 USA ayw3 D
0026 Pacific ayw2 D
2007 Pacific ayw2 D
3414 Pacific ayw2 D
5029 Pacific ayw2 D
A-16 Vietnam adw2 A
A-19 Japan adw2 A
I60 Italy adw2 A
I68 Italy adw2 A
I98 Italy adw2 A
7.32 Scotland adw2 A
B/19 Scotland adw2 A
B.21 Scotland adw2 A
B.31 Scotland adw2 A
9.5 Scotland adw2 A
U2 USA adw2 A
U4 USA adw2 A
U5 USA adw2 A
U6 USA adw2 A
U7 USA adw2 A
U9 USA adw2 A
U10 USA adw2 A
U8 USA ayw1 A
0012 India adw2 A

Appendix


Code Country Subtype/Genotype
Core Genes Amino Acid Sequences
110
120
130
140
150
160
170
180
A-4 Korea adw2 B LGFHISCVTFGRETALEYLVSLGGWIRTPTAY*PPNAPILSTLPETTGVRR--RGRSPRRRTPWPRRRRSQSPRRRRSQSRESQC*
A-34 China adw2 B .W.....L.....V.....F.V.....P..R.....V.....S.....*
A-38 China adw2 B .W.....L.....V.....F.V.....QP.R.....V.....S.....*
A-42 China adw2 B .W.....L.....V.....F.V.....P..R.....V.....S.....Y..*
A-43 China adw2 B .W.....L.....V.....F.V.....P..R.....V.....S.....*
A-45 China adw2 B .W.....L....D.V.....F.V.....P..R.Q.....V..G.....S.....P..*
A-48 China adw2 B .W.....L.....V.....F.V.....P..R.....V.....S.....*.*
A-11 Vietnam ayw1 B .W.....L.....V.....F.V.....P..R.Q.....V.....S.....P..*
A-12 Vietnam ayw1 B .W.....L.....V.....F.V.....P..R.Q.....V.....H.....S.....P..*
A-14 Vietnam ayw1 B .W.....L.....V.....F.V.....P..R.....V.....S.....R.*
A-15 Vietnam ayw1 B .W.....L.....V.....F.V.....P..R.Q.....V.....C.....S.....E.*
A-1 Korea adr C .W.....L.....V.....F.V.....PG.R.Q.....V..P.....S.....G...*
A-2 Korea adr C .W.....L.....V.....F.V.....P..R.....V.....S.....*
A-3 Korea adr C .W....SL.....G.....F.V.....Q..R.....S.....*
A-5 Korea adr C .W.....L.....V.....F.V.....P..R.Q.....V.....C.....S.....H.....*
A-8 Korea adr C .W.....L.....V.....F.V.....P..R.....V.....S.....*
A-9 Korea adr C .W.....L.....V.....F.V.....R.....V.....S.....*
A-10 Korea adr C .W.....L.....V.....F.V.....PG.R.....V.....S.....*
A-13 Vietnam adr C .W.....L.....V.....F.V.....P..R.....V.....S.....R.*
A-20 Japan adr C .W.....L.....V.....F.V.....R.....V.....S.....*
A-21 Japan adr C .W.....L.....V.....F.V.....P..R.....V..C.....S.....*
A-23 Japan adr C .W.....L.....V.....F.V.....P..R.....V.....S.....*
A-24 Japan adr C .W.....L.....V.....F.V.....S..R.....V.....S.....*
A-32 China adr C .W.....L.....V.....F.V.....R.....V.....S.....*
A-35 China adr C .W.....L.....V.....F.V.....P..R.....V.....S.....*
A-36 China adr C .W.....L.....V.....F.V.....P..R.....V.....S.....*
A-37 China adr C .W.....L.....V.....F.V.....P..R.....V.....S.....*
A-40 China adr C .W.....L.....V.....F.V.....P..R.....V..G.....S.....S.....P..*
A-41 China adr C .W.....L.....V.....F.V.....P..R.....V.....S.....Y..*
A-44 China adr C .W.....L.....V.....F.V.....P..R.Q.....V.....S.....*
A-49 China adr C .W.....L.....V.....F.V.....P..R.....V.....S.....P..*
A-50 China adr C .W.....L.....V.....F.V.....P..R.....V.....G.....S.....*
A-51 China adr C .W.....L.....V.....F.V.....P..R.....V.....S.....*
A-52 China adr C .W.....L.....V.....F.V.....P..R.Q.....V.....S.....P..*
A-53 China adr C .W.....L.....V.....F.V.....R.....V.....S.....K.....*
A-54 China adr C .W.....L.....V.....F.V.....P..R.....V.....C.....H.....S.....*
A-55 China adr C .W.....L.....V.....F.V.....P..R.....V.....S.....*
U3 USA adr C .W.....L.....V.G....F.V.....P..R.....V..G...C.....S.....N.*
0078 Pacific adrq- C .W.....L.....V.....F.V.....P..R.....V.....S.....*
2080 Pacific adrq- C .W.....L.....V.....F.V.....P..R.....V.....S.....*
3097 Pacific adrq- C .W.....L.....V.....F.V.....P..R.....V.....S.....*
3415 Pacific adrq- C .W.....L.....V.....F.V.....P..R.....V.....S.....*
5017 Pacific adrq- C .W.....L.....V.....F.V.....P..R.....V.....S.....*
5072 Pacific adrq- C .W.....L.....V.....F.V.....P..R.....V.....S.....*

Appendix

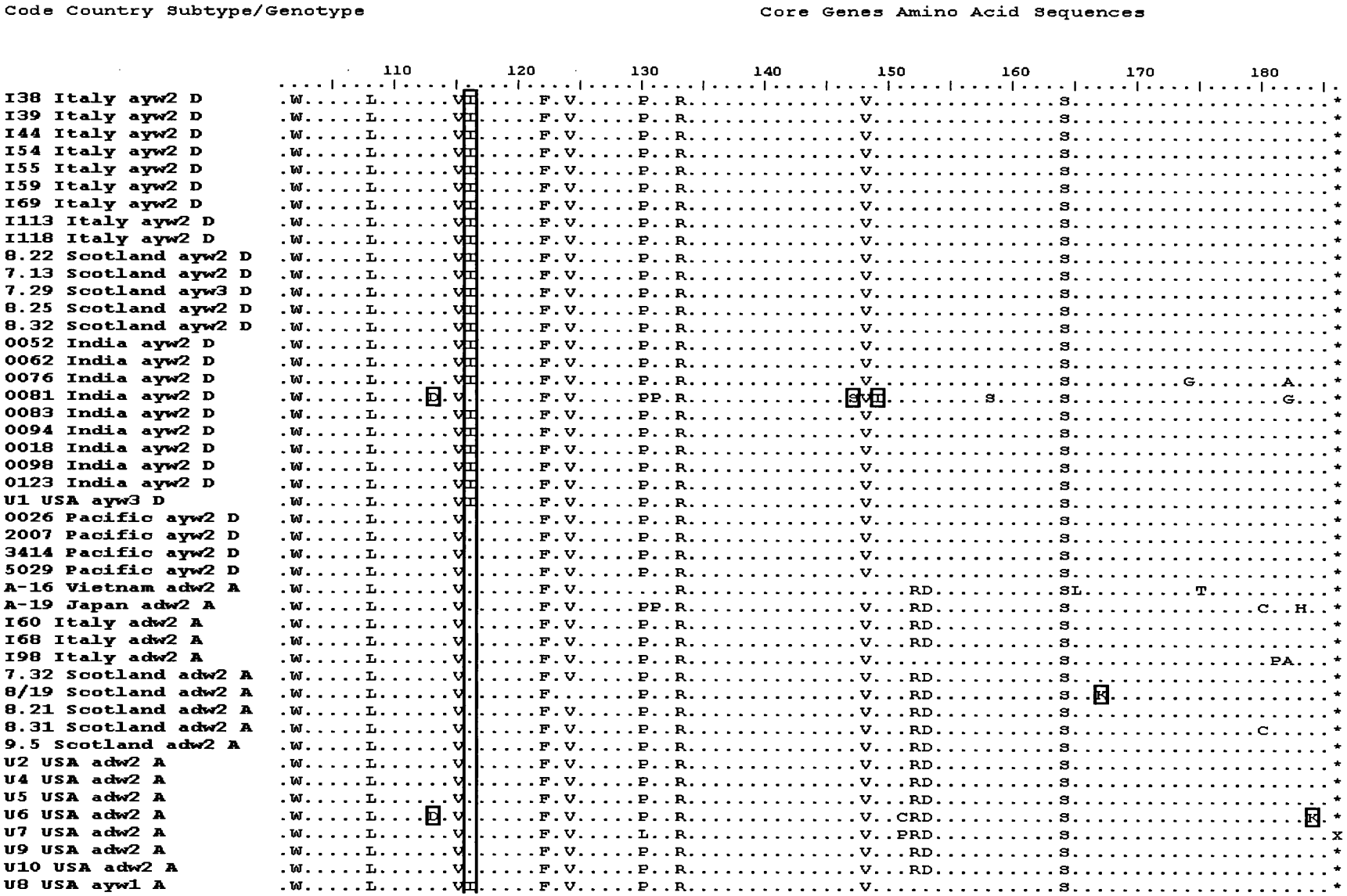
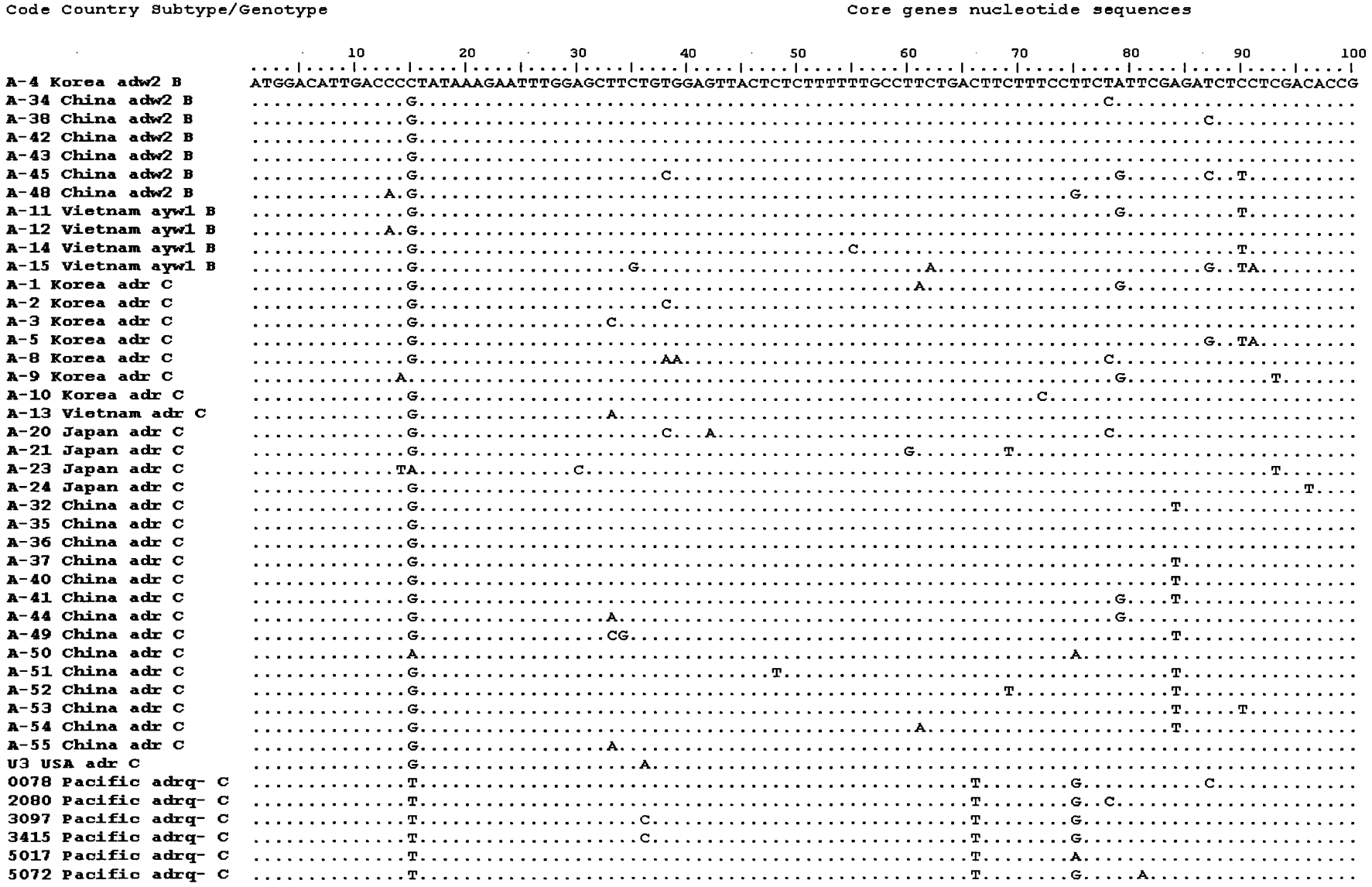

Code Country Subtype/Genotype
Core Genes Amino Acid Sequences
110
120
130
140
150
160
170
180
I38 Italy ayw2 D
I39 Italy ayw2 D
I44 Italy ayw2 D
I54 Italy ayw2 D
I55 Italy ayw2 D
I59 Italy ayw2 D
I69 Italy ayw2 D
I113 Italy ayw2 D
I118 Italy ayw2 D
8.22 Scotland ayw2 D
7.13 Scotland ayw2 D
7.29 Scotland ayw3 D
8.25 Scotland ayw2 D
8.32 Scotland ayw2 D
0052 India ayw2 D
0062 India ayw2 D
0076 India ayw2 D
0081 India ayw2 D
0083 India ayw2 D
0094 India ayw2 D
0018 India ayw2 D
0098 India ayw2 D
0123 India ayw2 D
U1 USA ayw3 D
0026 Pacific ayw2 D
2007 Pacific ayw2 D
3414 Pacific ayw2 D
5029 Pacific ayw2 D
A-16 Vietnam adw2 A
A-19 Japan adw2 A
I60 Italy adw2 A
I68 Italy adw2 A
I98 Italy adw2 A
7.32 Scotland adw2 A
8/19 Scotland adw2 A
8.21 Scotland adw2 A
8.31 Scotland adw2 A
9.5 Scotland adw2 A
U2 USA adw2 A
U4 USA adw2 A
U5 USA adw2 A
U6 USA adw2 A
U7 USA adw2 A
U9 USA adw2 A
U10 USA adw2 A
U8 USA ayw1 A

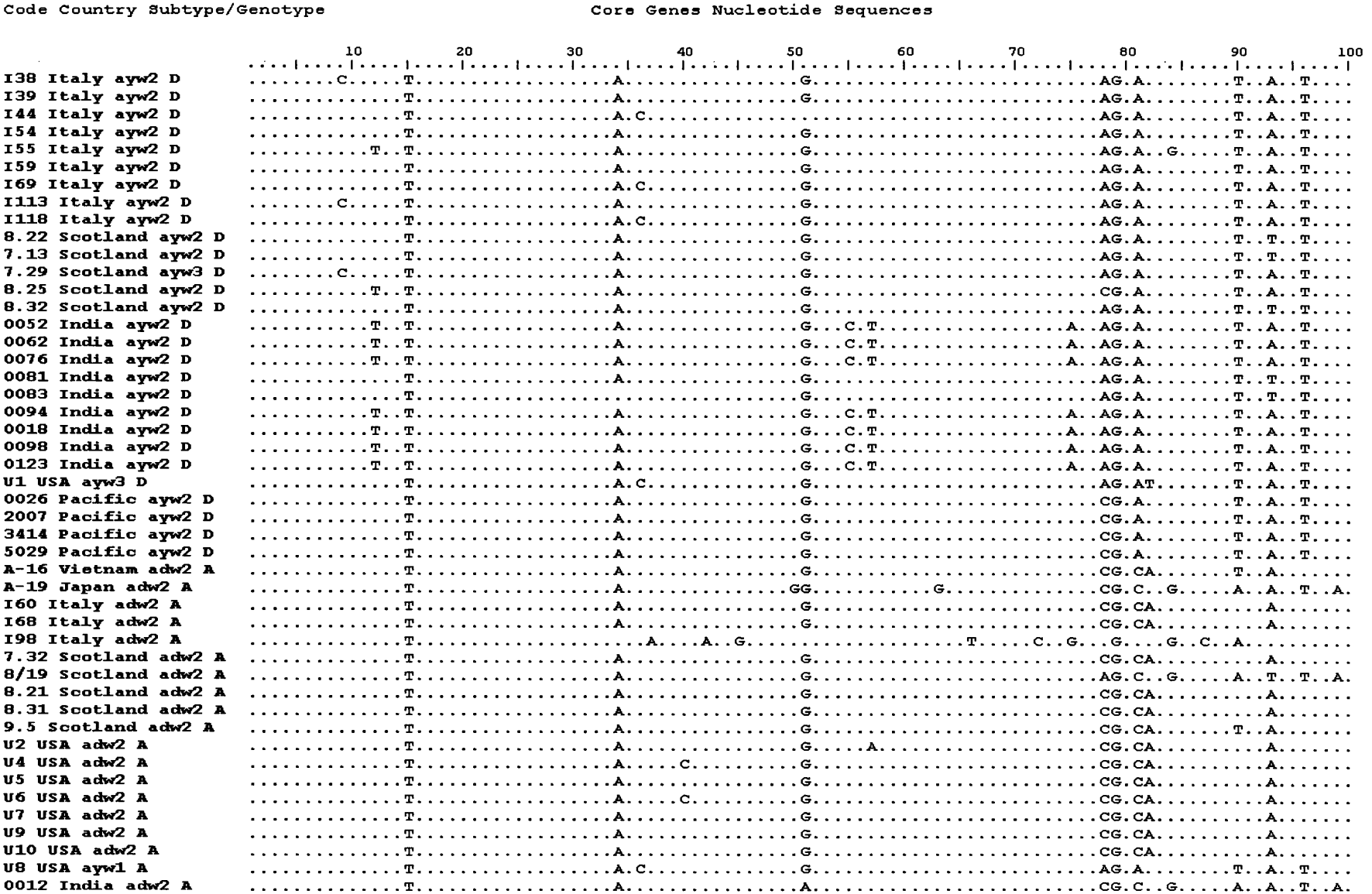
Table A.I.2- Alignment of complete nucleotide sequences of HBcAg which shows genotype/subtype identification and geographic origins of 91 sera.

Note: Nucleotides are numbered from the beginning of the HBcAg. Genotype A had 6 insertions, therefore using Bioedit software (Clustal W alignment) 6 extra nucleotides were added at positions 2454 to 2459 for a proper alignment.

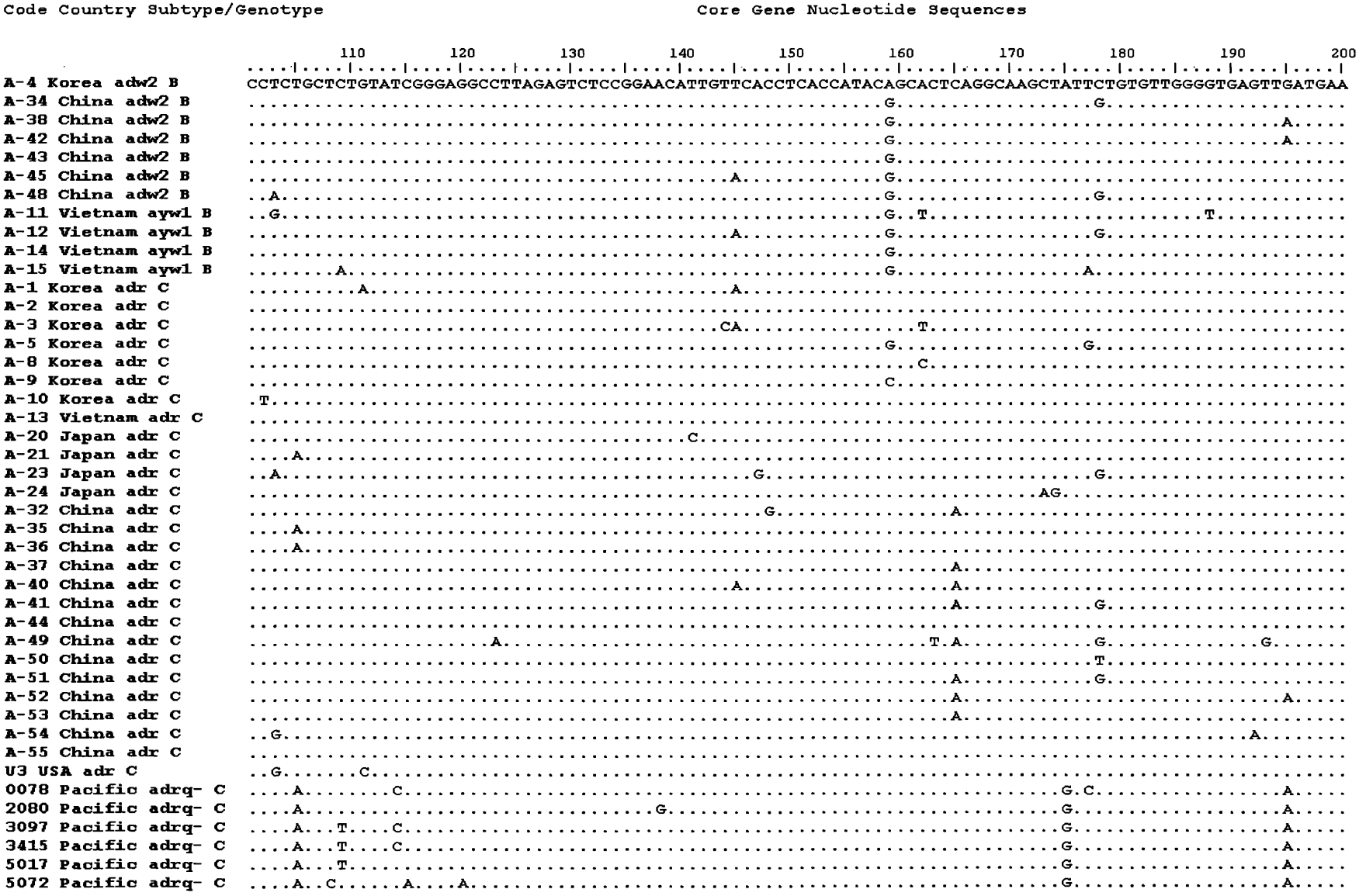
Appendix


Code Country Subtype/Genotype
Core genes nucleotide sequences
10 20 30 40 50 60 70 80 90 100
A-4 Korea adw2 B ATGGACATTGACCCCTATAAAGAATTTGGAGCTTCTGTGGAGTTACTCTCTTTTTTGCCTTCTGACTTCTTTCCTTCTATTCGAGATCTCCTCGACACCG
A-34 China adw2 B
A-38 China adw2 B
A-42 China adw2 B
A-43 China adw2 B
A-45 China adw2 B
A-48 China adw2 B
A-11 Vietnam ayw1 B
A-12 Vietnam ayw1 B
A-14 Vietnam ayw1 B
A-15 Vietnam ayw1 B
A-1 Korea adr C
A-2 Korea adr C
A-3 Korea adr C
A-5 Korea adr C
A-8 Korea adr C
A-9 Korea adr C
A-10 Korea adr C
A-13 Vietnam adr C
A-20 Japan adr C
A-21 Japan adr C
A-23 Japan adr C
A-24 Japan adr C
A-32 China adr C
A-35 China adr C
A-36 China adr C
A-37 China adr C
A-40 China adr C
A-41 China adr C
A-44 China adr C
A-49 China adr C
A-50 China adr C
A-51 China adr C
A-52 China adr C
A-53 China adr C
A-54 China adr C
A-55 China adr C
U3 USA adr C
0078 Pacific adrq- C
2080 Pacific adrq- C
3097 Pacific adrq- C
3415 Pacific adrq- C
5017 Pacific adrq- C
5072 Pacific adrq- C

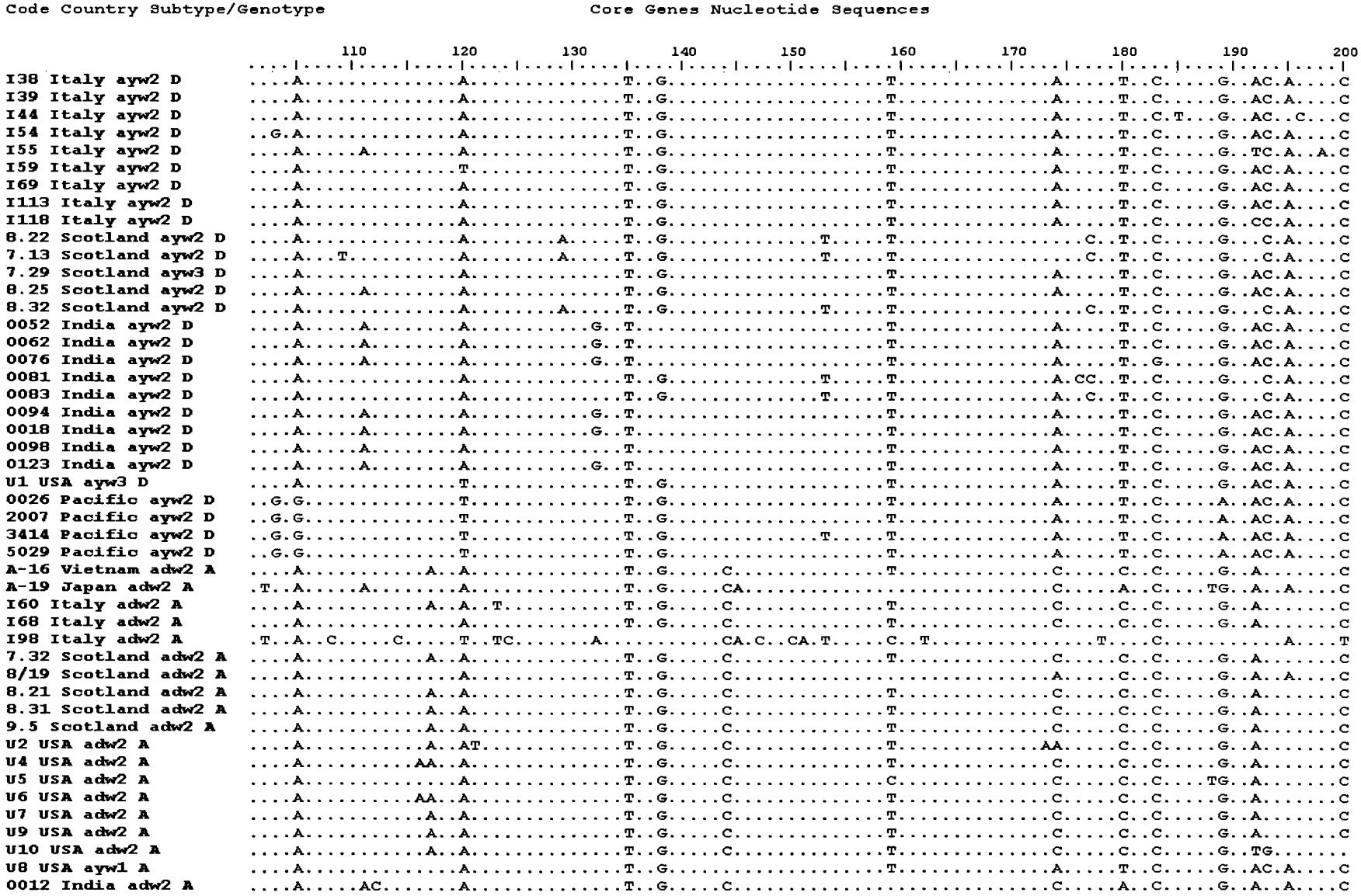
Appendix


Code Country Subtype/Genotype
Core Genes Nucleotide Sequences
10
20
30
40
50
60
70
80
90
100
I38 Italy ayw2 D
I39 Italy ayw2 D
I44 Italy ayw2 D
I54 Italy ayw2 D
I55 Italy ayw2 D
I59 Italy ayw2 D
I69 Italy ayw2 D
I113 Italy ayw2 D
I118 Italy ayw2 D
8.22 Scotland ayw2 D
7.13 Scotland ayw2 D
7.29 Scotland ayw3 D
8.25 Scotland ayw2 D
8.32 Scotland ayw2 D
0052 India ayw2 D
0062 India ayw2 D
0076 India ayw2 D
0081 India ayw2 D
0083 India ayw2 D
0094 India ayw2 D
0018 India ayw2 D
0098 India ayw2 D
0123 India ayw2 D
U1 USA ayw3 D
0026 Pacific ayw2 D
2007 Pacific ayw2 D
3414 Pacific ayw2 D
5029 Pacific ayw2 D
A-16 Vietnam adw2 A
A-19 Japan adw2 A
I60 Italy adw2 A
I68 Italy adw2 A
I98 Italy adw2 A
7.32 Scotland adw2 A
8/19 Scotland adw2 A
8.21 Scotland adw2 A
8.31 Scotland adw2 A
9.5 Scotland adw2 A
U2 USA adw2 A
U4 USA adw2 A
U5 USA adw2 A
U6 USA adw2 A
U7 USA adw2 A
U9 USA adw2 A
U10 USA adw2 A
U8 USA ayw1 A
0012 India adw2 A

Appendix


Code Country Subtype/Genotype
Core Gene Nucleotide Sequences
110 120 130 140 150 160 170 180 190 200
A-4 Korea adw2 B CCTCTGCTCTGTATCGGGAGGCCTTAGAGTCTCCGGAACATTGTTCACCTCACCATACAGCACTCAGGCAAGCTATTCTGTGTTGGGGTGAGTTGATGAA
A-34 China adw2 B
A-38 China adw2 B
A-42 China adw2 B
A-43 China adw2 B
A-45 China adw2 B
A-48 China adw2 B
A-11 Vietnam ayw1 B
A-12 Vietnam ayw1 B
A-14 Vietnam ayw1 B
A-15 Vietnam ayw1 B
A-1 Korea adr C
A-2 Korea adr C
A-3 Korea adr C
A-5 Korea adr C
A-8 Korea adr C
A-9 Korea adr C
A-10 Korea adr C
A-13 Vietnam adr C
A-20 Japan adr C
A-21 Japan adr C
A-23 Japan adr C
A-24 Japan adr C
A-32 China adr C
A-35 China adr C
A-36 China adr C
A-37 China adr C
A-40 China adr C
A-41 China adr C
A-44 China adr C
A-49 China adr C
A-50 China adr C
A-51 China adr C
A-52 China adr C
A-53 China adr C
A-54 China adr C
A-55 China adr C
U3 USA adr C
0078 Pacific adrq- C
2080 Pacific adrq- C
3097 Pacific adrq- C
3415 Pacific adrq- C
5017 Pacific adrq- C
5072 Pacific adrq- C

Appendix

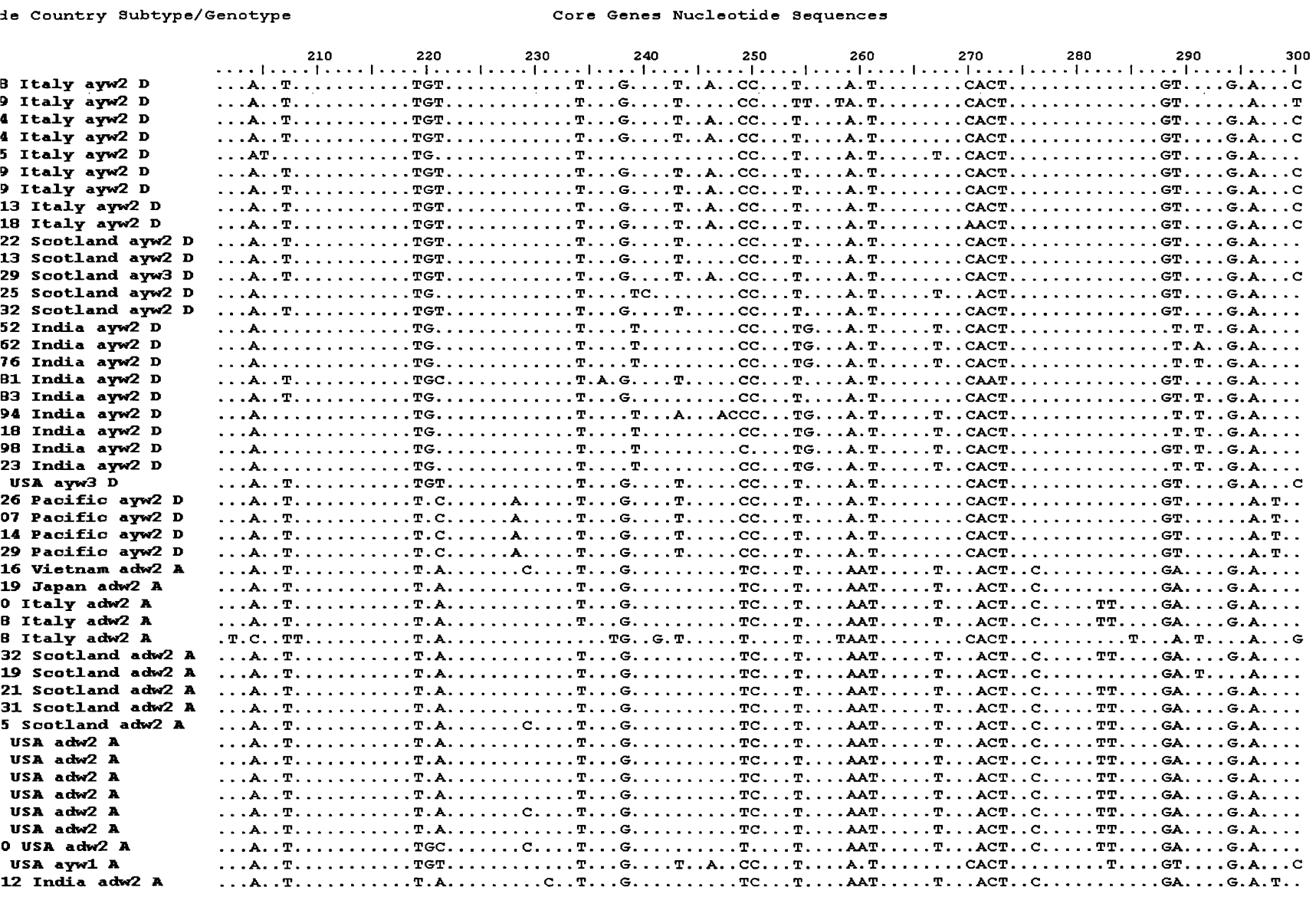

Code Country Subtype/Genotype
Core Genes Nucleotide Sequences
110
120
130
140
150
160
170
180
190
200
I38 Italy ayw2 D
I39 Italy ayw2 D
I44 Italy ayw2 D
I54 Italy ayw2 D
I55 Italy ayw2 D
I59 Italy ayw2 D
I69 Italy ayw2 D
I113 Italy ayw2 D
I118 Italy ayw2 D
8.22 Scotland ayw2 D
7.13 Scotland ayw2 D
7.29 Scotland ayw3 D
8.25 Scotland ayw2 D
8.32 Scotland ayw2 D
0052 India ayw2 D
0062 India ayw2 D
0076 India ayw2 D
0081 India ayw2 D
0083 India ayw2 D
0094 India ayw2 D
0018 India ayw2 D
0098 India ayw2 D
0123 India ayw2 D
U1 USA ayw3 D
0026 Pacific ayw2 D
2007 Pacific ayw2 D
3414 Pacific ayw2 D
5029 Pacific ayw2 D
A-16 Vietnam adw2 A
A-19 Japan adw2 A
I60 Italy adw2 A
I68 Italy adw2 A
I98 Italy adw2 A
7.32 Scotland adw2 A
8/19 Scotland adw2 A
8.21 Scotland adw2 A
8.31 Scotland adw2 A
9.5 Scotland adw2 A
U2 USA adw2 A
U4 USA adw2 A
U5 USA adw2 A
U6 USA adw2 A
U7 USA adw2 A
U9 USA adw2 A
U10 USA adw2 A
U8 USA ayw1 A
0012 India adw2 A

Appendix

Code Country Subtype/Genotype — Core Gene Nucleotide Sequences

```
                                 210       220       230       240       250       260       270       280       290       300
                                 ....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|
A-4 Korea adw2 B                 TCTGGCCACCTGGGTGGGAAGTAATTTGGAAGACCCAACATCCAGGGAATTAGCAGTCGGCTATGTCAATGTGAATATGGGCCTAAAACTCAGACCACTA
A-34 China adw2 B                ...A.....................T...G.....T....A.....C..T.....A.....A....
A-38 China adw2 B                ...A.....................T...G.....T.....C..T.....A....
A-42 China adw2 B                ...A.....................T...T.....T....A.....C..T.....A.....A....
A-43 China adw2 B                ...A.....................T...G.....T....A.....C..T.....A.T....A....
A-45 China adw2 B                ...A.....A.C.....G.....TG...AA.....T.....T..C.....T.....A.T..
A-48 China adw2 B                ...A.....................T...G.....T....A.....C..T.....A.....A....
A-11 Vietnam aywl B              ...A.....................T...G.....T....A.....T.....A.....A....
A-12 Vietnam aywl B              ...A.....................G.....T....A.....T.....A.....A....
A-14 Vietnam aywl B              .....................G.....T....A.....T.....A.....A.T..
A-15 Vietnam aywl B              .T.A.....................G.....T....A.....T.....A....
A-1 Korea adr C                  ...A.....G.....T...G.....GC..T....A.....C..T.....A.....A.....A....
A-2 Korea adr C                  .....................G.....T.....T.....A....
A-3 Korea adr C                  .....................G.....T.....T.....G.....A....
A-5 Korea adr C                  .T.A.....................G.....T....A.....T.....A....
A-8 Korea adr C                  .....................G.....T.....T.....A....
A-9 Korea adr C                  .....................G.....T....A.....T.....A.....A....
A-10 Korea adr C                 ...A.....................T...G.....T....A.....T.....A.....A....
A-13 Vietnam adr C               .....................G.....T....A.....T.....G..GA.....A....
A-20 Japan adr C                 .....................T...G.....T....A.....T.....A.....A....
A-21 Japan adr C                 C.....................G.....A.....T....A.....T.....A.T....A....
A-23 Japan adr C                 ......A.....................T...G.....G.T....A.....T.....A.....A....
A-24 Japan adr C                 .....................G.....T....A.....T.....A....
A-32 China adr C                 .....................G.....A.....T...AA.....T.....T.....A.A..
A-35 China adr C                 .....................G.....T....A.....T.....A.....A....
A-36 China adr C                 .....................G.....T....A.....T.....A.....A....
A-37 China adr C                 .....................GG.....T...AA.....T.....GA.....A....
A-40 China adr C                 .....................GG.T.....T...AA.....T....T.....A....
A-41 China adr C                 .....................GG.....G.T...AA.....T.....A.....A....
A-44 China adr C                 .....................T...G.....T....A.....T.....GA.....A....
A-49 China adr C                 .....................GG.....T...TA.....GT.....A.....A....
A-50 China adr C                 .....................G.....T....A.....T.....A.....A....
A-51 China adr C                 .....................T..GG.....T...AA.....T.....A.....A....
A-52 China adr C                 .....................GG.....T...AA.....C..T.....A...A..A....
A-53 China adr C                 .....C.....G.....C..T...AA.....C..T.....A.....A....
A-54 China adr C                 ......A.....TG.....T...AA.....C..T.....A.....A....
A-55 China adr C                 .....G.....T....A.....T.....GA.....A....
U3 USA adr C                     C.....CG.....T...G.....T.....T.....A.A..
0078 Pacific adrq- C             ......T.....G.....TG...A.T.....CA.T.....A.....A....
2080 Pacific adrq- C             ......T.....T...G.....TT...A.T.....C..T.....A.....A....
3097 Pacific adrq- C             ......T.....G.....TT...A.T.....T...A.T.....A.....A....
3415 Pacific adrq- C             ......T.....G.....TT...A.T.....T..CA.T.....A.....A....
5017 Pacific adrq- C             ......T.....G.....TT...A.T.....CA.T.....A.....A....
5072 Pacific adrq- C             ...A..T.....T.....CG.....T....A.T.....T.....A.....A.T..
```

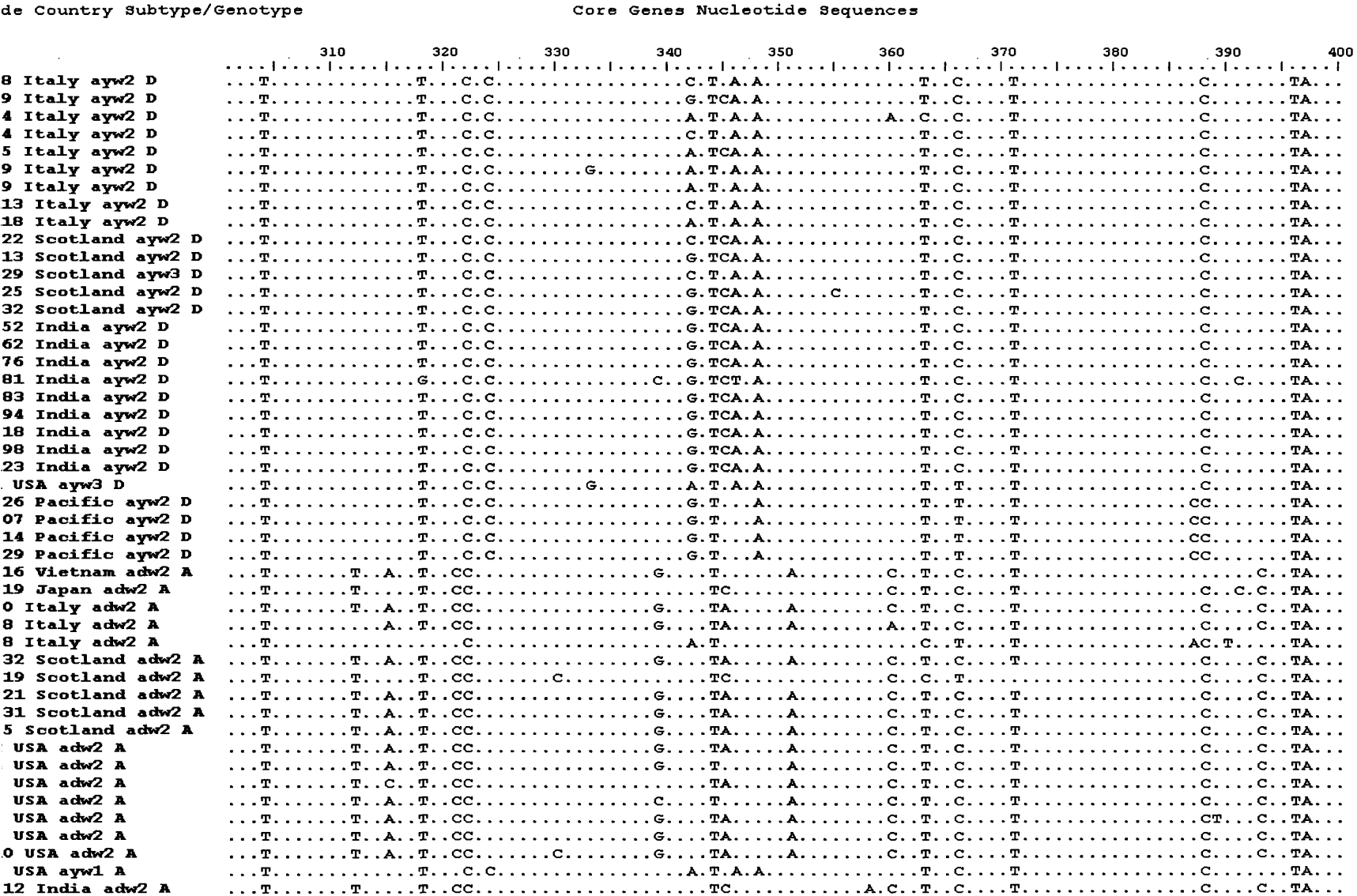
Appendix


de Country Subtype/Genotype
Core Genes Nucleotide Sequences
210
220
230
240
250
260
270
280
290
300
B Italy ayw2 D ...A..T.....TGT.....T...G....T..A..CC...T....A.T.....CACT.....GT....G.A...C
9 Italy ayw2 D ...A..T.....TGT.....T...G....T.....CC...TT..TA.T.....CACT.....GT.....A...T
4 Italy ayw2 D ...A..T.....TGT.....T...G....T..A..CC...T....A.T.....CACT.....GT....G.A...C
4 Italy ayw2 D ...A..T.....TGT.....T...G....T..A..CC...T....A.T.....CACT.....GT....G.A...C
5 Italy ayw2 D ...AT.....TG.....T.....CC...T....A.T.....T..CACT.....GT....G.A....
9 Italy ayw2 D ...A..T.....TGT.....T...G....T..A..CC...T....A.T.....CACT.....GT....G.A...C
9 Italy ayw2 D ...A..T.....TGT.....T...G....T..A..CC...T....A.T.....CACT.....GT....G.A...C
13 Italy ayw2 D ...A..T.....TGT.....T...G....T..A..CC...T....A.T.....CACT.....GT....G.A...C
18 Italy ayw2 D ...A..T.....TGT.....T...G....T..A..CC...T....A.T.....AACT.....GT....G.A...C
22 Scotland ayw2 D ...A..T.....TGT.....T...G....T.....CC...T....A.T.....CACT.....GT....G.A....
13 Scotland ayw2 D ...A..T.....TGT.....T...G....T.....CC...T....A.T.....CACT.....GT....G.A....
29 Scotland ayw3 D ...A..T.....TGT.....T...G....T..A..CC...T....A.T.....CACT.....GT....G.A...C
25 Scotland ayw2 D ...A.....TG.....T....TC.....CC...T....A.T.....T...ACT.....GT....G.A....
32 Scotland ayw2 D ...A..T.....TGT.....T...G....T.....CC...T....A.T.....CACT.....GT....G.A....
52 India ayw2 D ...A.....TG.....T....T.....CC...TG...A.T.....T..CACT.....T.T..G.A....
62 India ayw2 D ...A.....TG.....T....T.....CC...TG...A.T.....T..CACT.....T.A..G.A....
76 India ayw2 D ...A.....TG.....T....T.....CC...TG...A.T.....T..CACT.....T.T..G.A....
B1 India ayw2 D ...A..T.....TGC.....T.A.G....T.....CC...T....A.T.....CAAT.....GT....G.A....
B3 India ayw2 D ...A..T.....TG.....T...G.....CC...T....A.T.....CACT.....GT.T..G.A....
94 India ayw2 D ...A.....TG.....T....T...A...ACCC...TG...A.T.....T..CACT.....T.T..G.A....
18 India ayw2 D ...A.....TG.....T....T.....CC...TG...A.T.....T..CACT.....T.T..G.A....
98 India ayw2 D ...A.....TG.....T....T.....C....TG...A.T.....T..CACT.....GT.T..G.A....
23 India ayw2 D ...A.....TG.....T....T.....CC...TG...A.T.....T..CACT.....T.T..G.A....
USA ayw3 D ...A..T.....TGT.....T...G....T.....CC...T....A.T.....CACT.....GT....G.A...C
26 Pacific ayw2 D ...A..T.....T.C.....A.....T...G....T.....CC...T....A.T.....CACT.....GT.....A.T..
07 Pacific ayw2 D ...A..T.....T.C.....A.....T...G....T.....CC...T....A.T.....CACT.....GT.....A.T..
14 Pacific ayw2 D ...A..T.....T.C.....A.....T...G....T.....CC...T....A.T.....CACT.....GT.....A.T..
29 Pacific ayw2 D ...A..T.....T.C.....A.....T...G....T.....CC...T....A.T.....CACT.....GT.....A.T..
16 Vietnam adw2 A ...A..T.....T.A.....C....T...G.....TC...T....AAT.....T...ACT..C.....GA....G.A....
19 Japan adw2 A ...A..T.....T.A.....T...G.....TC...T....AAT.....T...ACT..C.....GA....G.A....
0 Italy adw2 A ...A..T.....T.A.....T...G.....TC...T....AAT.....T...ACT..C.....TT....GA....G.A....
8 Italy adw2 A ...A..T.....T.A.....T...G.....TC...T....AAT.....T...ACT..C.....TT....GA....G.A....
8 Italy adw2 A .T.C..TT.....T.A.....TG..G.T.....T....T...TAAT.....CACT.....T...A.T....A...G
32 Scotland adw2 A ...A..T.....T.A.....T...G.....TC...T....AAT.....T...ACT..C.....TT....GA....G.A....
19 Scotland adw2 A ...A..T.....T.A.....T...G.....TC...T....AAT.....T...ACT..C.....GA.T....A....
21 Scotland adw2 A ...A..T.....T.A.....T...G.....TC...T....AAT.....T...ACT..C.....TT....GA....G.A....
31 Scotland adw2 A ...A..T.....T.A.....T...G.....TC...T....AAT.....T...ACT..C.....TT....GA....G.A....
5 Scotland adw2 A ...A..T.....T.A.....C....T...G.....TC...T....AAT.....T...ACT..C.....TT....GA....G.A....
USA adw2 A ...A..T.....T.A.....T...G.....TC...T....AAT.....T...ACT..C.....TT....GA....G.A....
USA adw2 A ...A..T.....T.A.....T...G.....TC...T....AAT.....T...ACT..C.....TT....GA....G.A....
USA adw2 A ...A..T.....T.A.....T...G.....TC...T....AAT.....T...ACT..C.....TT....GA....G.A....
USA adw2 A ...A..T.....T.A.....T...G.....TC...T....AAT.....T...ACT..C.....TT....GA....G.A....
USA adw2 A ...A..T.....T.A.....C....T...G.....TC...T....AAT.....T...ACT..C.....TT....GA....G.A....
USA adw2 A ...A..T.....T.A.....T...G.....TC...T....AAT.....T...ACT..C.....TT....GA....G.A....
0 USA adw2 A ...A..T.....TGC.....C....T...G.....T....T....AAT.....T...ACT..C.....TT....GA....G.A....
USA ayw1 A ...A..T.....TGT.....T...G....T..A..CC...T....A.T.....CACT.....T....GT....G.A...C
12 India adw2 A ...A..T.....T.A.....C..T...G.....TC...T....AAT.....T...ACT..C.....GA....G.A.T..

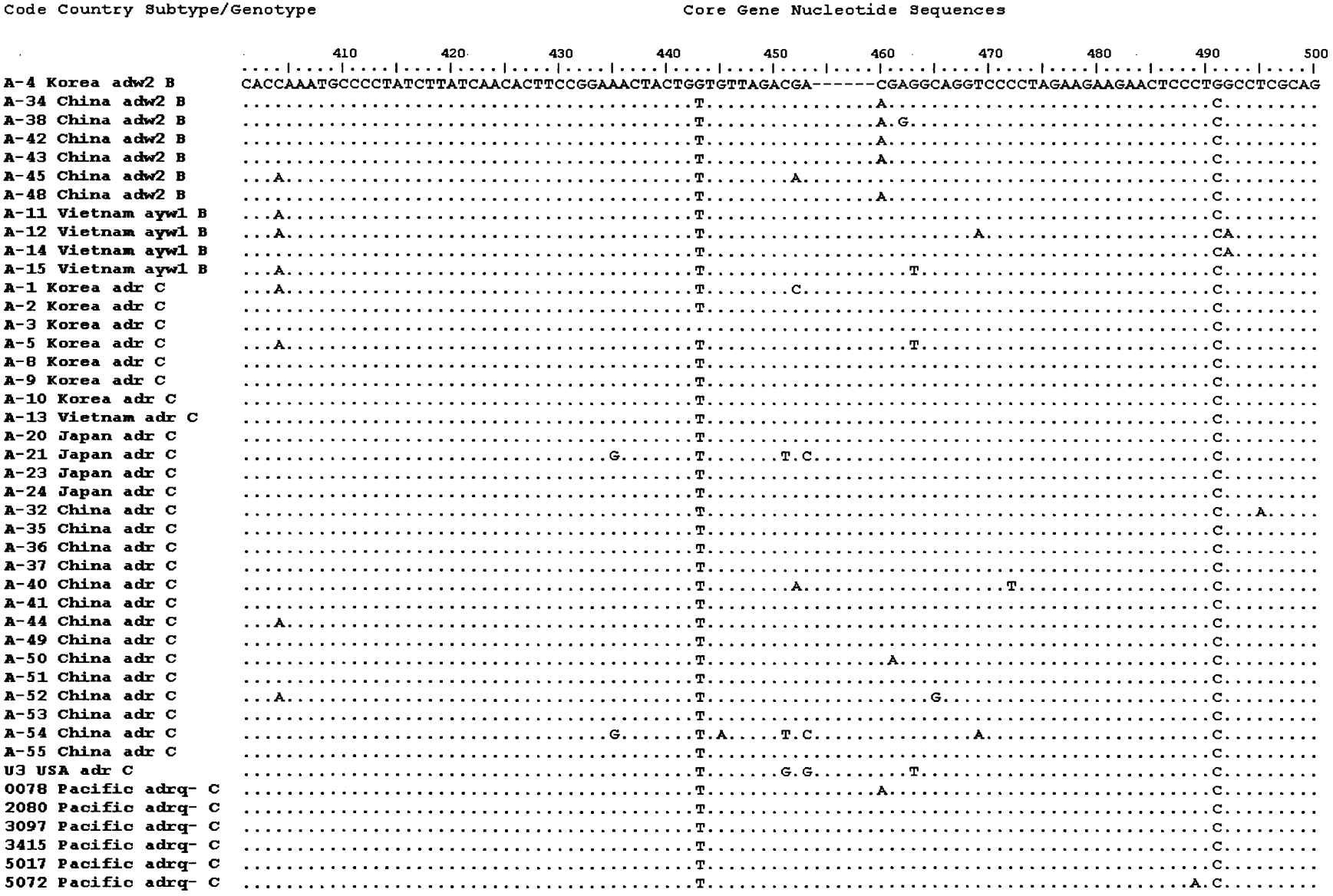
Appendix

Code	Country	Subtype/Genotype	Core Gene Nucleotide Sequences (310–400)
A-4	Korea	adw2 B	TTGGGGTTTCACATTTCCTGTGTTACTTTTGGAAGAGAAACTGCTCTTGAGTATTTGGTGTCATTAGGAGGGTGGATTCGCACTCCTACAGCTTACTGAC
A-34	China	adw2 B	...T.....C.....G.....T.....A.....T..T....T.....C.T..A..TA...
A-38	China	adw2 B	...T.....C.....G.....T.....A.....T..T....T.....CAGC.A..TA...
A-42	China	adw2 B	...T.....C.....G.....T.....A.....T..T....T.....C.T..A..TA...
A-43	China	adw2 B	...T.....C.....G.....T.....A.....T..T....T.....C.T..A..TA...
A-45	China	adw2 B	...T.....C.....T....T.A....A.....T..T....T.....C.T..A..TA...
A-48	China	adw2 B	...T.....C.....G.....T.....A...C....A..T..T....T.....C.T..A..TA...
A-11	Vietnam	ayw1 B	...T.....T...C.....T.....A.....T..T....T.....C.T..A..TA...
A-12	Vietnam	ayw1 B	...T.....C.....T.....A.....T..T....T.....C.T..A..TA...
A-14	Vietnam	ayw1 B	...T.....C.....T.....A.....T..T....T.....C.T..A..TA...
A-15	Vietnam	ayw1 B	...T.....C.....T.....A.....T..T....T.....C.T..A..TA...
A-1	Korea	adr C	...T.....T...C.....T.....T..T....T.....C.C.G....A...
A-2	Korea	adr C	...T.....C.....T.....T..T....T.....C.C.....A...
A-3	Korea	adr C	...T.....C.C.....G.....T..T....T.....CA.....A...
A-5	Korea	adr C	...T.....C.....T.....A.....T..T....T.....C.T..A..TA...
A-8	Korea	adr C	...T.....C.....T.....T..T....T.....C.C.....A...
A-9	Korea	adr C	...T.....C.....TG.....T..T....T.....C.....A...
A-10	Korea	adr C	...T.....C.....T.....A..T..C....T.....C.C.G....A...
A-13	Vietnam	adr C	...T.....C.....G.....T.....A..T..T....T.....C.C.....A...
A-20	Japan	adr C	...T.....C.C.....T.....T..T....T.....C..A...A...
A-21	Japan	adr C	C..T.....C.....T.....T..T....T.....C.T.....A...
A-23	Japan	adr C	C..T.....C.....C.....T.....T..T....T.....C.T.....A...
A-24	Japan	adr C	...T.....C.....T.....T..T....T.....T.C.....A...
A-32	China	adr C	...T.....C.....C.....T.....T..T....T.....C.....A...
A-35	China	adr C	...T.....C.....T.....T..T....T.....C.T.....A...
A-36	China	adr C	...T.....C.....T.....T..T....T.....C.T.....A...
A-37	China	adr C	...T.....C.....T.....T..T....T.....C.C.....A...
A-40	China	adr C	...T.....C.....T.....T..T....T.....C.C.....A...
A-41	China	adr C	...T.....C.....T.....T..T....T.....C.C.....A...
A-44	China	adr C	...T.....C.....G.....T.....A..T..T....T.....C.C.....A...
A-49	China	adr C	...T.....C.....T.....T..T....T.....C.T.....A...
A-50	China	adr C	...T.....C.....T.....T..T....T.....C.C.....A...
A-51	China	adr C	...T.....C.....T.....T..T....T.....C.C.....A...
A-52	China	adr C	...T.....CC.....T.....T..T....T.....C.C.....A...
A-53	China	adr C	...T.....C.....T.....T..T....T.....T.....A...
A-54	China	adr C	...T....C.....C.....T.....T..T....T.....C.C.....A...
A-55	China	adr C	...T.....C.....G.....T.....A..T..T....T.....C.C.....A...
U3	USA	adr C	...T....C.....C....G.....T....C.....T..T....T.....C.C.....TA...
0078	Pacific	adrq- C	C..T.....C.....T.....T..T....T.....C.T.....A...
2080	Pacific	adrq- C	...T.....C.....T.....T..T....T.....C.C.....A...
3097	Pacific	adrq- C	...T.....C.....T.....T..T....T.....C.C.....A...
3415	Pacific	adrq- C	...T.....C.....T.....T..T....T.....C.C.....A...
5017	Pacific	adrq- C	...T.....C.....T.....T..T....T.....C.C.....A...
5072	Pacific	adrq- C	...T.....C.....T.....T..T....T.....C.C..A...A...

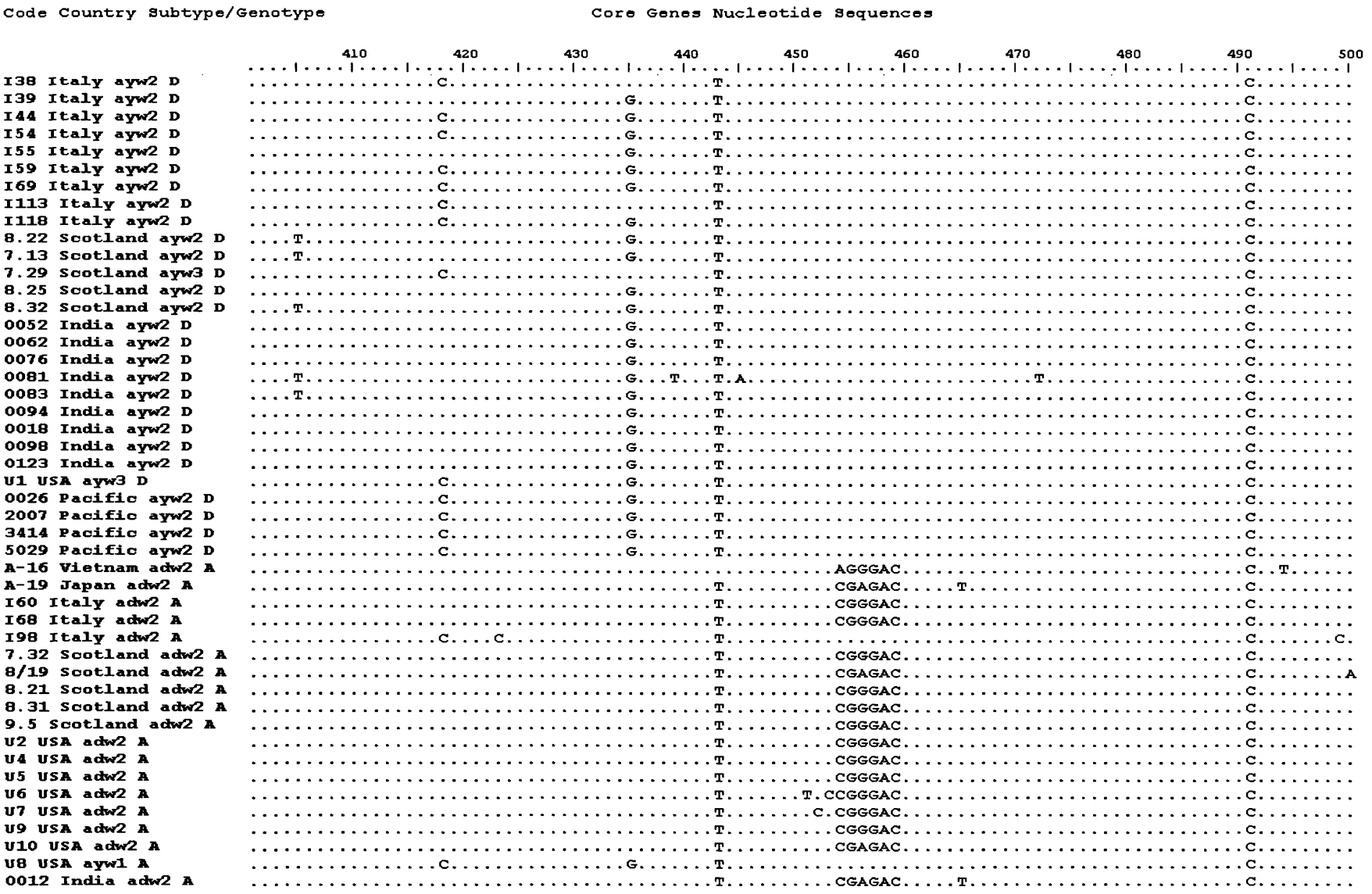
Appendix


de Country Subtype/Genotype
Core Genes Nucleotide Sequences
310
320
330
340
350
360
370
380
390
400
8 Italy ayw2 D
9 Italy ayw2 D
4 Italy ayw2 D
4 Italy ayw2 D
5 Italy ayw2 D
9 Italy ayw2 D
9 Italy ayw2 D
13 Italy ayw2 D
18 Italy ayw2 D
22 Scotland ayw2 D
13 Scotland ayw2 D
29 Scotland ayw3 D
25 Scotland ayw2 D
32 Scotland ayw2 D
52 India ayw2 D
62 India ayw2 D
76 India ayw2 D
81 India ayw2 D
83 India ayw2 D
94 India ayw2 D
18 India ayw2 D
98 India ayw2 D
23 India ayw2 D
USA ayw3 D
26 Pacific ayw2 D
07 Pacific ayw2 D
14 Pacific ayw2 D
29 Pacific ayw2 D
16 Vietnam adw2 A
19 Japan adw2 A
0 Italy adw2 A
8 Italy adw2 A
8 Italy adw2 A
32 Scotland adw2 A
19 Scotland adw2 A
21 Scotland adw2 A
31 Scotland adw2 A
5 Scotland adw2 A
USA adw2 A
USA adw2 A
USA adw2 A
USA adw2 A
USA adw2 A
USA adw2 A
0 USA adw2 A
USA ayw1 A
12 India adw2 A

Appendix

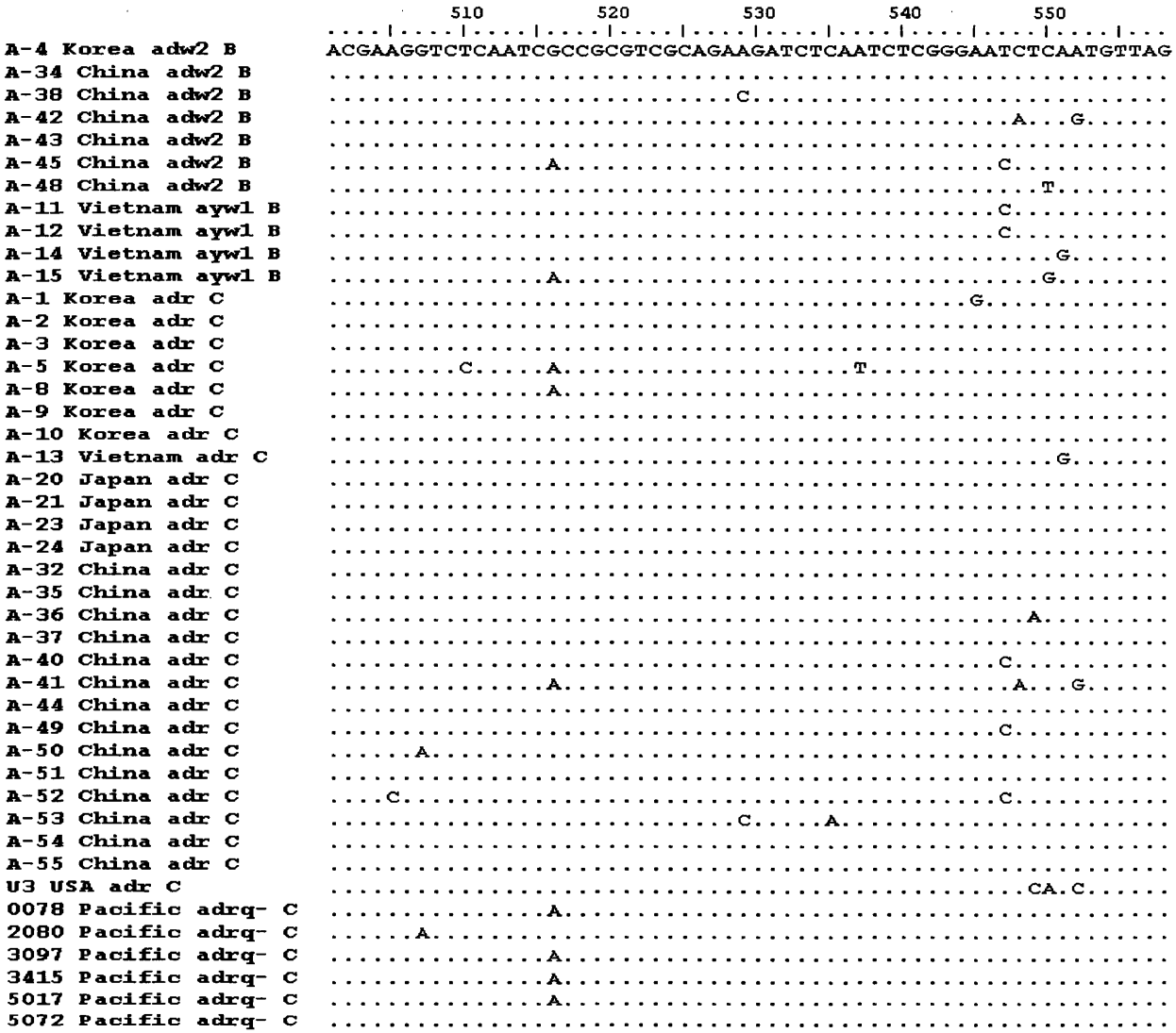

Code Country Subtype/Genotype
Core Gene Nucleotide Sequences
410
420
430
440
450
460
470
480
490
500
A-4 Korea adw2 B CACCAAATGCCCCTATCTTATCAACACTTCCGGAAACTACTGGTGTTAGACGA------CGAGGCAGGTCCCCTAGAAGAAGAACTCCCTGGCCTCGCAG
A-34 China adw2 B
A-38 China adw2 B
A-42 China adw2 B
A-43 China adw2 B
A-45 China adw2 B
A-48 China adw2 B
A-11 Vietnam ayw1 B
A-12 Vietnam ayw1 B
A-14 Vietnam ayw1 B
A-15 Vietnam ayw1 B
A-1 Korea adr C
A-2 Korea adr C
A-3 Korea adr C
A-5 Korea adr C
A-8 Korea adr C
A-9 Korea adr C
A-10 Korea adr C
A-13 Vietnam adr C
A-20 Japan adr C
A-21 Japan adr C
A-23 Japan adr C
A-24 Japan adr C
A-32 China adr C
A-35 China adr C
A-36 China adr C
A-37 China adr C
A-40 China adr C
A-41 China adr C
A-44 China adr C
A-49 China adr C
A-50 China adr C
A-51 China adr C
A-52 China adr C
A-53 China adr C
A-54 China adr C
A-55 China adr C
U3 USA adr C
0078 Pacific adrq- C
2080 Pacific adrq- C
3097 Pacific adrq- C
3415 Pacific adrq- C
5017 Pacific adrq- C
5072 Pacific adrq- C

Appendix


Code Country Subtype/Genotype
Core Genes Nucleotide Sequences
410
420
430
440
450
460
470
480
490
500
I38 Italy ayw2 D
I39 Italy ayw2 D
I44 Italy ayw2 D
I54 Italy ayw2 D
I55 Italy ayw2 D
I59 Italy ayw2 D
I69 Italy ayw2 D
I113 Italy ayw2 D
I118 Italy ayw2 D
8.22 Scotland ayw2 D
7.13 Scotland ayw2 D
7.29 Scotland ayw3 D
8.25 Scotland ayw2 D
8.32 Scotland ayw2 D
0052 India ayw2 D
0062 India ayw2 D
0076 India ayw2 D
0081 India ayw2 D
0083 India ayw2 D
0094 India ayw2 D
0018 India ayw2 D
0098 India ayw2 D
0123 India ayw2 D
U1 USA ayw3 D
0026 Pacific ayw2 D
2007 Pacific ayw2 D
3414 Pacific ayw2 D
5029 Pacific ayw2 D
A-16 Vietnam adw2 A
A-19 Japan adw2 A
I60 Italy adw2 A
I68 Italy adw2 A
I98 Italy adw2 A
7.32 Scotland adw2 A
8/19 Scotland adw2 A
8.21 Scotland adw2 A
8.31 Scotland adw2 A
9.5 Scotland adw2 A
U2 USA adw2 A
U4 USA adw2 A
U5 USA adw2 A
U6 USA adw2 A
U7 USA adw2 A
U9 USA adw2 A
U10 USA adw2 A
U8 USA ayw1 A
0012 India adw2 A

Appendix

Code Country Subtype/Genotype — Core Gene Nucleotide Sequences



Appendix

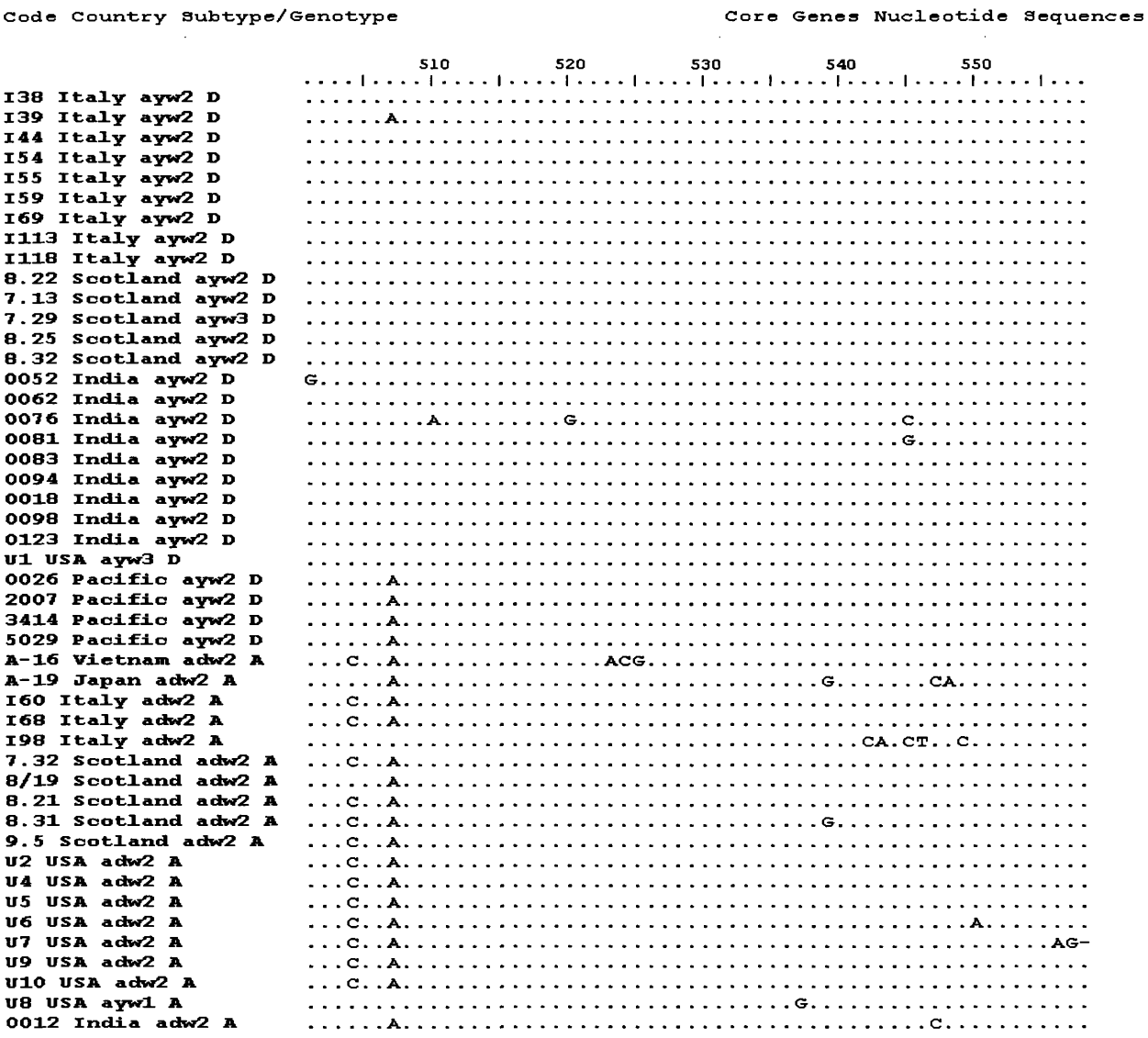

Code Country Subtype/Genotype
Core Genes Nucleotide Sequences
510
520
530
540
550
I38 Italy ayw2 D
I39 Italy ayw2 D
I44 Italy ayw2 D
I54 Italy ayw2 D
I55 Italy ayw2 D
I59 Italy ayw2 D
I69 Italy ayw2 D
I113 Italy ayw2 D
I118 Italy ayw2 D
8.22 Scotland ayw2 D
7.13 Scotland ayw2 D
7.29 Scotland ayw3 D
8.25 Scotland ayw2 D
8.32 Scotland ayw2 D
0052 India ayw2 D
0062 India ayw2 D
0076 India ayw2 D
0081 India ayw2 D
0083 India ayw2 D
0094 India ayw2 D
0018 India ayw2 D
0098 India ayw2 D
0123 India ayw2 D
U1 USA ayw3 D
0026 Pacific ayw2 D
2007 Pacific ayw2 D
3414 Pacific ayw2 D
5029 Pacific ayw2 D
A-16 Vietnam adw2 A
A-19 Japan adw2 A
I60 Italy adw2 A
I68 Italy adw2 A
I98 Italy adw2 A
7.32 Scotland adw2 A
8/19 Scotland adw2 A
8.21 Scotland adw2 A
8.31 Scotland adw2 A
9.5 Scotland adw2 A
U2 USA adw2 A
U4 USA adw2 A
U5 USA adw2 A
U6 USA adw2 A
U7 USA adw2 A
U9 USA adw2 A
U10 USA adw2 A
U8 USA ayw1 A
0012 India adw2 A

Table A.I.3) Alignment of complete HBcAg amino acid sequences from Fiii.

Code Core gene amino acid sequence from Fiji

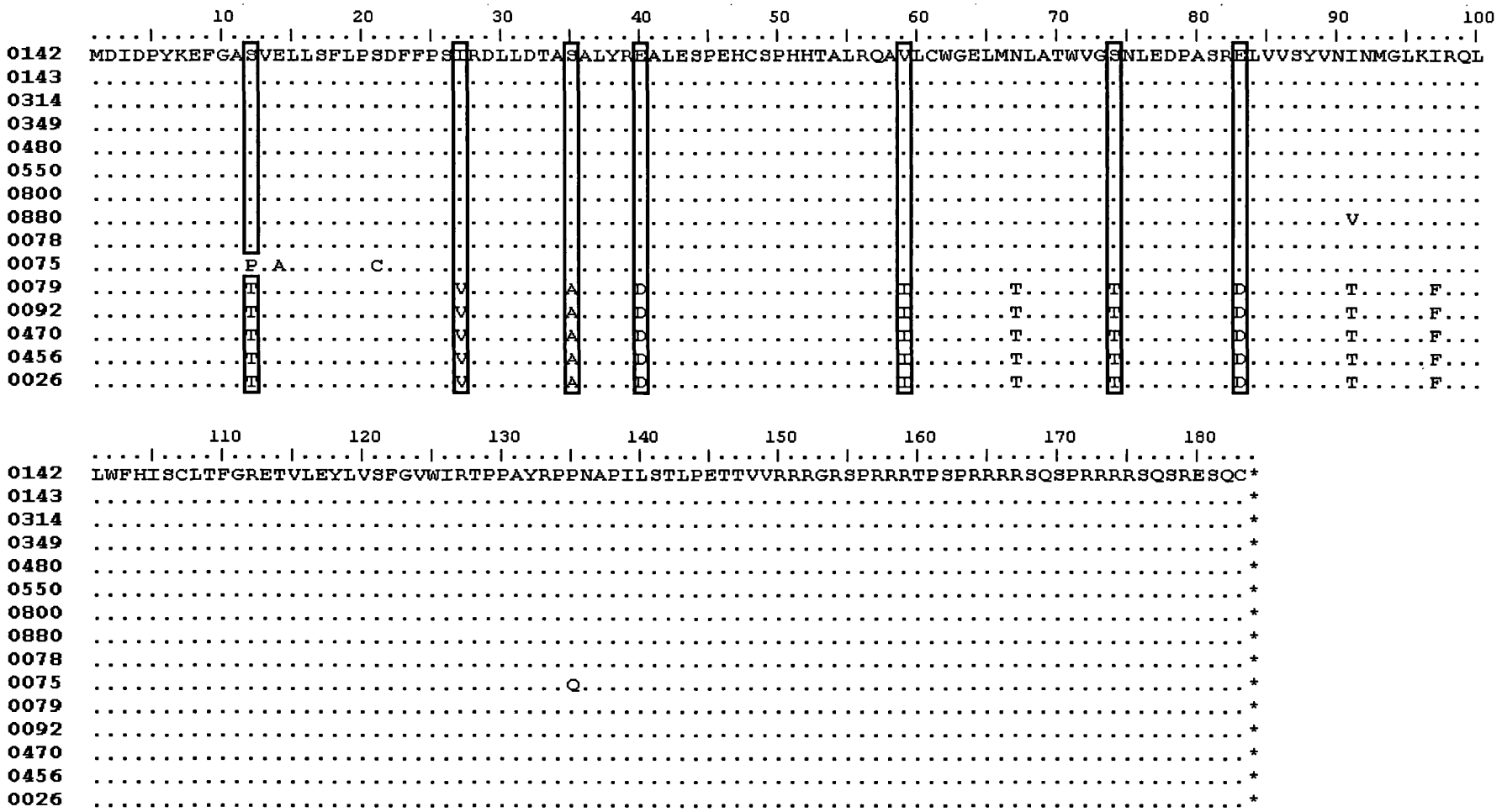


Table A.I.4) Alignment of complete nucleotide sequences from Fiji.

```
Code                                  Core gene nucleotide sequence from Fiji

                  10        20        30        40        50        60        70        80        90        100
                  |....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|
0142  ATGGACATTG ACCCTTATAA AGAATTTGGA GCTTCTGTGG AGTTACTCTC TTTTTTGCCT TCTGATTTCT TCCCGTCTAT CCGAGATCTC CTCGACACCG
0143  .......... .......... .......... .......... .......... .......... .......... .......... .......... ..........
0314  .......... .......... .......... .....C.... .......... .......... .......... .T........ T........T ..........
0349  .......... .......... .......... .......... .......... .......... .......... .......... T......... ..........
0480  .......... .......... .......... .......... .......... .......... .......... .T........ T......... ..........
0550  .......... ....A..... .......... .......... .......... .......... .......... .T........ T......... ..........
0800  .......... .......... .......... .......... .......... .......... .......... .T........ T......... ..........
0880  .......... .......... .......... .......... .......... .......... .......... .T.....C.. T..G.....T ..........
0078  .......... .......... .......... .......... .......... .......... .......... .T........ T.....C... ..........
0075  .......... .......... .......... ...C.C.... C......... .......... .G........ .T........ T......... ..........
0079  .......... .......... .......... ...A...... .......... G......... .....C.... .T..T..CG. A........T ..A..T....
0092  .......... .......... .......... ...A...... .......... G......... .....C.... .T..T..CG. A........T ..A..T....
0470  .......... .......... .......... ...A...... .......... G......... .....C.... .T..T..CG. A........T ..A..T....
0456  .......... .......... .......... ...A.C.... .......... G......... .....C.... .T..T..CG. A........T ..A..T....
0026  .......... .......... .......... ...A...... .......... G......... .....C.... .T..T..CG. A........T ..A..T....

                  110       120       130       140       150       160       170       180       190       200
                  |....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|
0142  CCTCAGCTCT GTATCGGGAG GCCTTAGAGT CTCCGGAACA TTGTTCACCT CACCATACAG CACTCAGGCA AGCTGTTCTG TGTTGGGGTG AGTTAATGAA
0143  .......... .......... .......... .......... .......... .......... .......... .......... .......... ..........
0314  ........T. ...C...... .......... .......... .......... .......... .......... .......... .......... ..........
0349  .......... .......... .......... .......... .......... .......... .......... .......... .......... ..........
0480  .......... .......... .......... .......... .......... .......... .......... .......... .......... ..........
0550  .......... A..C...... .......... .......... .......... .......... .......... .......... .......... ..........
0800  .......... .......... .......... .......... .......... .......... .......... .......... .......... ..........
0880  .......... .......... .......... .......G.. .......... .......... .......... .......... .......... ..........
0078  .......... ...C...... .......... .......... .......... .......... .......... ......C... .......... ..........
0075  ........T. ...C...... .......... .......... .......... .......... .......... .......... .......... ..........
0079  ..G.G..... .........T .......... ....T..G.. .......... ........T. .......... ...AA....T ..C.....A. .AC......C
0092  ..G.G..... .........T .......... ....T..G.. .......... ........T. .......... ...AA....T ..C.....A. .AC......C
0470  ..G.G..... .........T ..T..G.... ....T..G.. .......... ........T. .......... ...AA....T ..C.....A. .AC......C
0456  ..G.G..... .........T .......... ....T..G.. .......... ........T. .......... ...AA....T ..C.....A. .AC......C
0026  ..G.G..... .........T .......... ....T..G.. .......... ........T. .......... ...AA....T ..C.....A. .AC......C
```

Appendix

Code — Core gene nucleotide sequence from Fiji

```
          210        220        230        240        250        260        270        280        290        300
     ....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....|
0142 TCTGGCTACC TGGGTGGGAA GTAATTTGGA AGACCCAGCA TCCCGGGAAT TAGTTGTCAG CTATGTCAAC ATTAATATGG GCCTAAAAAT CAGACAACTA
0143 .......... .......... .......... .......... .......... .......... .......... .......... .......... ..........
0314 .......... .......... .......... .......... ...A...... .......... T.....T... .......... .......... ..........
0349 .......... .......... .......... .......... ...A...... ........A. .......... .......... .......... ..........
0480 .......... .......... .......... .......... ...A...... .......... T......... .......... .......... ..........
0550 .......... .......... .......... .......... ...A...... ....G..... T......... .......... .......... ..........
0800 .......... .......... .......... .......... ...A...... .......... T......... .......... .......... ..........
0880 .......... .......... .......... .......... ...A.A.... ....A..... T......... G......... .......... ..........
0078 .......... .......... .......... .......... ...A...... ....G..... T......... .......... .......... ..........
0075 .......... .......... .......... .......... ...A...... .......... T.....T... .......... .......... ..........
0079 ...A...... ........T. C......A.. ...T...... ..TA....CC ....A..... T......... .C........ .......GT. .......T..
0092 ...A...... ........T. C......A.. ...T...... ..TA....CC ....A..... T......... .C........ .......GT. .......T..
0470 ...A...... ........T. C......A.. ...T...... ..TA....CC ....A..... T......... .C........ .......GT. .......T..
0456 ...A...... ........T. C......A.. ...T...... ..TA....CC ....A..... T......... .C........ .......GT. T......T..
0026 ...A...... ........T. C......A.. ...T...... ..TA....CC ....A..... T......... .C........ .......GT. .......T..

          310        320        330        340        350        360        370        380        390        400
     ....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....|
0142 TTGTGGTTTC ACATTTCCTG TCTCACTTTT GGAAGAGAAA CTGTTCTTGA GTATTTGGTG TCTTTTGGAG TGTGGATTCG CACTCCTCCC GCTTACAGAC
0143 .......... .......... .......... .......... .......... .......... .......... .......... .........A ..........
0314 .......... .......... ...T...... .......... .......... .......... .......... .......... .......... ..........
0349 .......... .......... .......... .......... .......... .......... .......... .......... .......... ..........
0480 .......... .......... ...T...... .......... .......... .......... .......... .......... .......... ..........
0550 .......... .......... ...T...... .......... .......... .......... .......... .......... .......... ..........
0800 .......... .......... ...T...... .......... .......... ...C...... .......... .......... .......... ..........
0880 .......... .......... ...T...... .......... .......... ...C...... .....C.... .......... .......... ..........
0078 C......... .......... ...T...... .......... .......... .......... .......... .......... .........T ..........
0075 .......... .......... ...T...... .......... ........G. .......... .......... .......... .......... ..........
0079 .......... .......T.. .......... .......... .G.....A.. .......... .......... .......... ......C..A .....T....
0092 .......... .......T.. .......... .......... .G.....A.. .......... .......... .......... ......C..A .....T....
0470 .......... .......T.. .......... .......... .G.....A.. .......... .......... .......... ......C..A .....T....
0456 .......... .......T.. .......... .......... .G.....A.. .......... .......... .......... ......C..A .....T....
0026 .......... .......T.. .......... .......... .G.....A.. .......... .......... .......... ......C..A .....T....
```

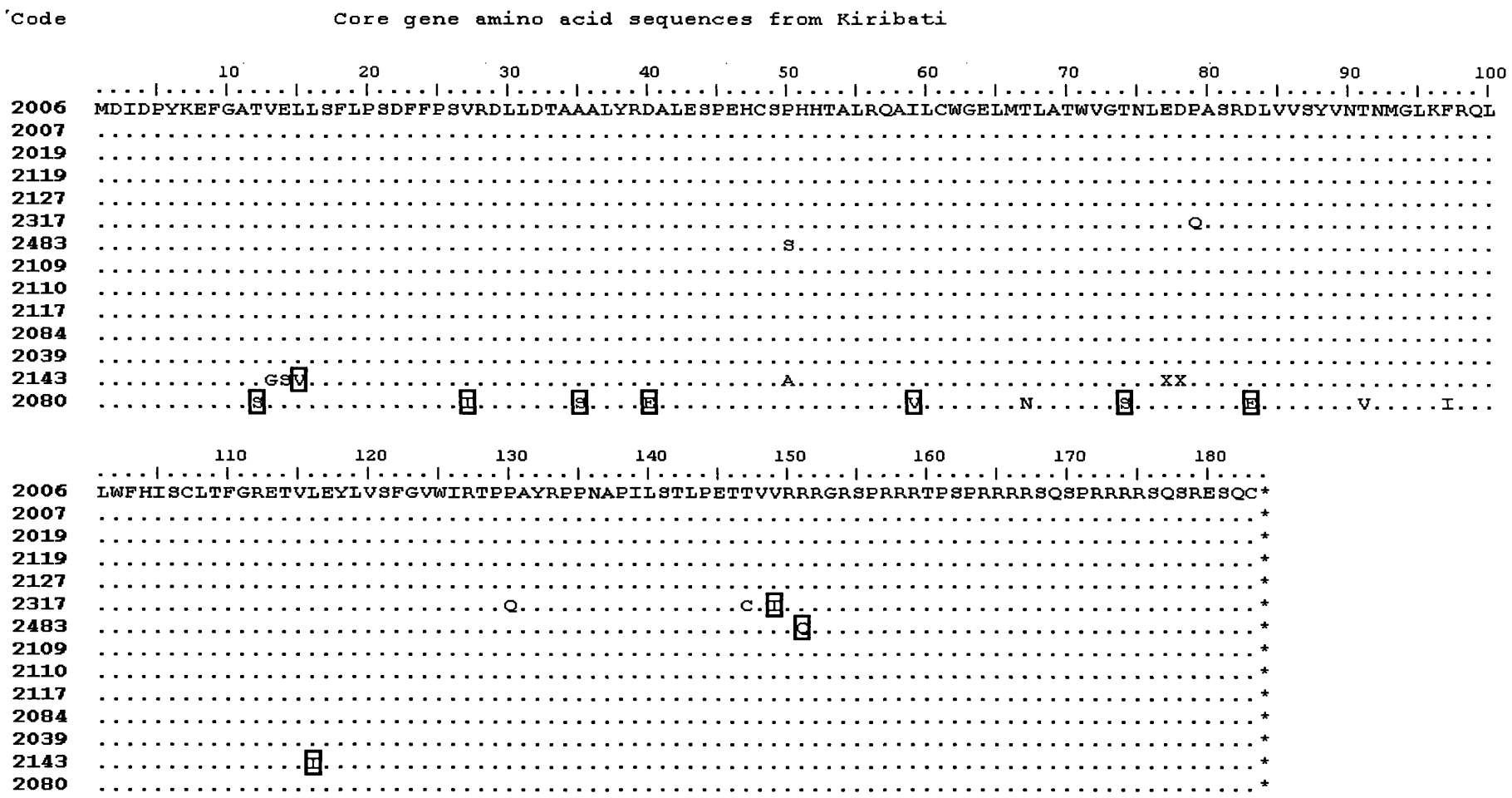
Appendix

Code Core gene nucleotide sequence from Fiji

```
          410        420        430        440        450        460        470        480        490        500
     ....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....|
0142 CACCAAATGC CCCTATCCTA TCAACACTTC CGGAAACTAC TGTTGTTAGA CGACGAGGCA GGTCCCCTAG AAGAAGAACT CCCTCGCCTC GCAGACGAAG
0143 .......... .......... .......... .......... .......... .......... .......... .......... .......... ..........
0314 .......... .......T.. .......... .......... .......... .......... .......... .......... .......... ..........
0349 .......... .......... .......... .......... .......... .......... .......... .......... .......... ..........
0480 .......... .......T.. .......... .......... .......... .......... .......... .......... .......... ..........
0550 .......... .......T.. .......... .......... .......... .......... .......... .......... .......... ..........
0800 .......... .......T.. .......... .......... .......... .......... .......... .......... .......... ..........
0880 .......... .......T.. .......... .......... .......... ...A...... .......... .......... .......... ..........
0078 .......... .......T.. .......... .......... .......... ...A...... .......... .......... .......... ..........
0075 ...A...... .......T.. .......... .......... .......... .......... .......... .......... .......... ..........
0079 .......... .......... .......... ....G..... .......... .......... .......... .......... .......... ..........
0092 .......... .......... .......... ....G..... .......... .......... .......... .......... .......... ..........
0470 .......... .......... .......... ....G..... .......... .......... .......... .......... .......... ..........
0456 .......... .......... .......... ....G..C.. .......... .......... .......... .......... .......... ..........
0026 .......... .......... .......... ....G..... .......... .......... .......... .......... .......... ..........

          510        520        530        540        550
     ....|....| ....|....| ....|....| ....|....| ....|....| ..
0142 ATCTCAATCA CCGCGTCGCA GAAGATCTCA ATCTCGGGAA TCTCAATGTT AG
0143 .........G .......... .......... .......... .......... ..
0314 G......... .......... .......... .......... .......... ..
0349 .......... .......... .......... .......... .......... ..
0480 G......... .......... .......... .......... .......... ..
0550 G......... .......... .......... .......... .......... ..
0800 G......... .......... .......... .......... .......... ..
0880 .........G .......... .......... .......... .......... ..
0078 G......... .......... .......... .......... .......... ..
0075 G......... .......... .......... .......... .......... ..
0079 .........G .......... .......... .......... .......... ..
0092 .........G .......... .......... .......... .......... ..
0470 .........G .......... .......... .......... .......... ..
0456 .........G .......... .......... .......... .......... ..
0026 .........G .......... .......... .......... .......... ..
```

Appendix


Code
Core gene amino acid sequences from Kiribati
10 20 30 40 50 60 70 80 90 100
2006 MDIDPYKEFGATVELLSFLPSDFFPSVRDLLDTAAALYRDALESPEHCSPHHTALRQAILCWGELMTLATWVGTNLEDPASRDLVVSYVNTNMGLKFRQL
2007
2019
2119
2127
2317
2483
2109
2110
2117
2084
2039
2143
2080
110 120 130 140 150 160 170 180
2006 LWFHISCLTFGRETVLEYLVSFGVWIRTPPAYRPPNAPILSTLPETTVVRRRGRSPRRRTPSPRRRRSQSPRRRRSQSRESQC*
2007
2019
2119
2127
2317
2483
2109
2110
2117
2084
2039
2143
2080

Appendix

Table A.I.6) Alignment of complete nucleotide sequences from Kiribati.

Code — Core gene nucleotide sequences from Kiribati

```
          10        20        30        40        50        60        70        80        90        100
          |....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|
2006      ATGGACATTGACCCTTATAAAGAATTTGGAGCTACTGTGGAGTTACTCTCGTTTTTGCCTTCTGACTTCTTTCCTTCCGTACGAGATCTTCTAGATACCG
2007      ................................................................................................
2019      ................................................................................................
2119      ................................................................................................
2127      ................................................................................................
2317      ..............G.................................................................................
2483      ..........................................................T.........T.................C....
2109      ................................................................................................
2110      ................................................................................................
2117      ................................................................................................
2084      ................................................................................................
2039      ................................................................................................
2143      ....................................G.AGTG......................................................
2080      ................................T...............T...............T.........G...A.T.........C..C..C....

          110       120       130       140       150       160       170       180       190       200
          |....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|
2006      CCGCGGCTCTGTATCGGGATGCCTTAGAGTCTCCTGAGCATTGTTCACCTCACCATACTGCACTCAGGCAAGCAATTCTTTGCTGGGGAGAACTAATGAC
2007      ................................................................................................
2019      ................................................................................................
2119      ................................................................................................
2127      ................................................................................................
2317      ................................................................................................
2483      ............................A...................T.....................................................
2109      ................................................................................................
2110      ................................................................................................
2117      ................................................................................................
2084      ..............................................................C.................................
2039      ................................................................................................
2143      ..........................................G.....................................................
2080      ..T.A..............G..............G.......................A..............TG....G..T.....T..GT......A
```

Appendix

Code Core gene nucleotide sequences from Kiribati

```
          210       220       230       240       250       260       270       280       290       300
          ....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|
2006      TCTAGCTACCTGGGTGGGTACTAATTTAGAAGATCCAGCATCTAGGGACCTAGTAGTCAGTTATGTCAACACTAATATGGGCCTAAAGTTCAGACAATTA
2007      ..............................................................................................
2019      ......C.......................................................................................
2119      ..............................................................................................
2127      ..............................................................................................
2317      ..................................A...........................................................
2483      ..............................................................................................
2109      ......C.......................................................................................
2110      ..............................................A...............................................
2117      .T............................................................................................
2084      ..............................................................................................
2039      ..............................................................................................
2143      ..................C.........---.....................................................C........
2080      ...G.............A.G......G.............C.....AT....T.............GT.............AA........C..

          310       320       330       340       350       360       370       380       390       400
          ....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|
2006      TTGTGGTTTCACATTTCTTGTCTCACTTTTGGAAGAGAAACGGTTCTAGAGTATTTGGTGTCTTTTGGAGTGTGGATTCGCACTCCCCCAGCTTATAGAC
2007      ..............................................................................................
2019      ..............................................................................................
2119      ..............................................................................................
2127      ..............................................................................................
2317      ..............................................................................................
2483      ..............................................................................................
2109      ..............................................................................................
2110      ..............................................................................................
2117      ..........................................T...................................T...............
2084      ..............................................................................................
2039      ..............................................................................................
2143      ..........................................A...................................................
2080      ..................C.....T.................T.....T.....................................T..C.....C....
```

Appendix

Code Core gene nucleotide sequences from Kiribati

```
          410        420        430        440        450        460        470        480        490        500
     ....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|
2006 CACCAAATGCCCCTATCCTATCAACACTTCCGGAGACTACTGTTGTTAGACGACGAGGCAGGTCCCCTAGAAGAAGAACTCCCTCGCCTCGCAGACGAAG
2007 ..................................................................................................
2019 ..................................................................................................
2119 ..................................................................................................
2127 ..................................................................................................
2317 ........................................TG....A.....................................................
2483 ..............................................A...................................................
2109 ..................................................................................................
2110 ..................................................................................................
2117 ..................................................................................................
2084 ..................................................................................................
2039 ..................................................................................................
2143 ...................................A..............................................................
2080 ..................T...............A..............................................................

          510        520        530        540        550
     ....|....|....|....|....|....|....|....|....|....|..
2006 ATCTCAATCGCCGCGTCGCAGAAGATCTCAATCTCGGGAATCTCAATGTTAG
2007 ...................................................
2019 ...................................................
2119 ...................................................
2127 ...................................................
2317 ...................................................
2483 ...................................................
2109 ...................................................
2110 ...................................................
2117 ...................................................
2084 ...................................................
2039 ...................................................
2143 G........A..........................................
2080 ...................................................
```

Table A.I.7) Alignment of complete amino acid sequences from Tonga.

```
Code                Core gene amino acid sequences from Tonga

               10        20        30        40        50        60        70        80        90       100
      ....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|
3415  MDIDPYKEFGASVELLSFLPSDFFPSIRDLLDTASALYREALESPEHCSPHHTALRQAVLCWGELMNLATWVGSNLEDPASRELVVSYVNINMGLKIRQL
3417  ....................................................................................................
3097  ....................................................................................................
3419  ....................................................................................................
3428  ....................................................................................................
3629  ....................................................................................................
3099  ....................................................................................................
3221  ....................................................................................................
3365  ....................................................................................................
3369  ....................................................................................................
3519  ....................................................................................................
3309  ............................................................W............RS.D..T.M..................
3509  ..................................................................D.................................
3343  ...........T..............V.......A....D...................I......T......T........D.......T.....F...
3414  ...........T..............V.......A....D...................I......T......T........D.......T.....F...

              110       120       130       140       150       160       170       180
      ....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....
3415  LWFHISCLTFGRETVLEYLVSFGVWIRTPPAYRPPNAPILSTLPETTVVRRRGRSPRRRTPSPRRRRSQSPRRRRSQSRESQC*
3417  ...................................................................................*
3097  ...................................................................................*
3419  ...................................................................................*
3428  ...................................................................................*
3629  ...................G...................................................P...........*
3099  ...................................................................................*
3221  ...................................................................................*
3365  ...................................................................................*
3369  ...................................................................................*
3519  ...................................................................................*
3309  .............................T.....................................................*
3509  ..................................F...........C....................................*
3343  ............................................A......................................*
3414  ...................................................................................*
```

Table A.I.8) Alignment of complete nucleotide sequences from Tonga.

```
Code                 Core nucleotide sequences from Tonga

                   10        20        30        40        50        60        70        80        90       100
          ....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|
3415      ATGGACATTGACCCTTATAAAGAATTTGGAGCTTCCGTGGAGTTACTCTCTTTTTTGCCTTCTGATTTCTTTCCGTCTATTCGAGATCTCCTCGACACCG
3417      ................................C...................................................................
3097      ....................................................................................................
3419      ....................................................................................................
3428      ....................................................................................................
3629      .........................................A..........................................................
3099      ...................................T.........................................C......................
3221      ...................................T................................................................
3365      ...................................T..................................................C.............
3369      ...................................T..................................................C.............
3519      ...................................T..................................................C.............
3309      ...................................T................................................................
3509      ...................................T...................................C............................
3343      .................................A.T..............G..............C........T..CG.A........T..A..T....
3414      .................................A.T..............G..............C........T..CG.A........T..A..T....

                  110       120       130       140       150       160       170       180       190       200
          ....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|
3415      CCTCAGCTTTGTACCGGGAGGCCTTAGAGTCTCCGGAACATTGTTCACCTCACCATACAGCACTCAGGCAAGCTGTTCTGTGTTGGGGTGAGTTAATGAA
3417      .............................................................................T......................
3097      ....................................................................................................
3419      ....................................................................................................
3428      ....................................................................................................
3629      .....................................G..............................................................
3099      ........C...........................................................................................
3221      ........C....T......................................................................................
3365      ........C...........................................................................................
3369      ........C...........................................................................................
3519      ........C...........................................................................................
3309      ........C.....................................G...................................G.................
3509      ........C..........................................................................................G.
3343      ..G.G...C....T.....T..............T..G....................T..............AA....T..C.....A..AC......C
3414      ..G.G...C....T.....T..............T..G..............T.....T..............AA....T..C.....A..AC......C
```

Appendix

Code Core nucleotide sequences from Tonga

```
          210       220       230       240       250       260       270       280       290       300
     ....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|
3415 TCTGGCTACCTGGGTGGGAAGTAATTTGGAAGACCCAGCATCCAGGGAATTAGTTGTCAGTTATGTTAACATTAATATGGGCCTAAAAATCAGACAACTA
3417 ..............................................................................................
3097 ..........................................................T...................................
3419 ..............................................................................................
3428 ..............................................................................................
3629 ..............................................................................................
3099 ..................................................G...........C.................................
3221 ..................................................G...........C.................................
3365 ..................................................G...........C.................................
3369 ..................................................G...........C.................................
3519 ..................................................G...........C.................................
3309 ....................GTC......C......A......T.........G...........C.................................
3509 C...............................T.....T.....A........G.....C.....C.................................
3343 ...A..............T.C......A.....T........T.....CC....A...........C....C..............GT........T..
3414 ...A..............T.C......A.....T........T.....CC....A...........C....C..............GT........T..

          310       320       330       340       350       360       370       380       390       400
     ....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|
3415 TTGTGGTTTCACATTTCCTGTCTTACTTTTGGAAGAGAAACTGTTCTTGAGTATTTGGTGTCTTTTGGAGTGTGGATTCGCACTCCTCCCGCTTACAGAC
3417 C.............................................................................................
3097 ..............................................................................................
3419 ..................................................A...........................................
3428 ..............................................................................................
3629 ..........................................................GT..................................
3099 ..............................................................................................
3221 C.............................................................................................
3365 ..............................................................................C...............
3369 ..............................................................................C...............
3519 ..............................................................................C...............
3309 ..............................................................................................A.............
3509 ..............................................................................................
3343 ..................T.....C..................G.....A.....................................C..A.....T....
3414 ..................T.....C..................G.....A.....................................C..A.....T....
```

Appendix

Code Core nucleotide sequences from Tonga

```
          410        420        430        440        450        460        470        480        490        500
     ....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|
3415 CACCAAATGCCCCTATCTTATCAACACTTCCGGAAACTACTGTTGTTAGACGACGAGGCAGGTCCCCTAGAAGAAGAACTCCCTCGCCTCGCAGACGAAG
3417 ..................................................................................................
3097 ..................................................................................................
3419 ..................................................................................................
3428 ..................................................................................................
3629 ..................................................................................................A....
3099 ..................................................................................................A....
3221 ..................................................................................................
3365 ..................................................................................................
3369 ..................................................................................................
3519 ..................................................................................................
3309 ..................................................................................................
3509 ..TTC.....................................TG..................................................
3343 ................C.............CG..................................................
3414 ................C.............G...................................................

          510        520        530        540        550
     ....|....|....|....|....|....|....|....|....|....|..
3415 GTCTCAATCACCGCGTCGCAGAAGATCTCAATCTCGGGAATCTCAATGTTAG
3417 ...................................................
3097 ...................................................
3419 ...................................................
3428 ...................................................
3629 .........C....C.....................................
3099 ...................................................
3221 ...................................................
3365 ...................................................
3369 ...................................................
3519 ...................................................
3309 ...................................................
3509 A..................................................
3343 A........G..........................................
3414 A........G..........................................
```

Appendix

Table A.I.9) Alignment of complete amino acid sequences from Vanuatu.

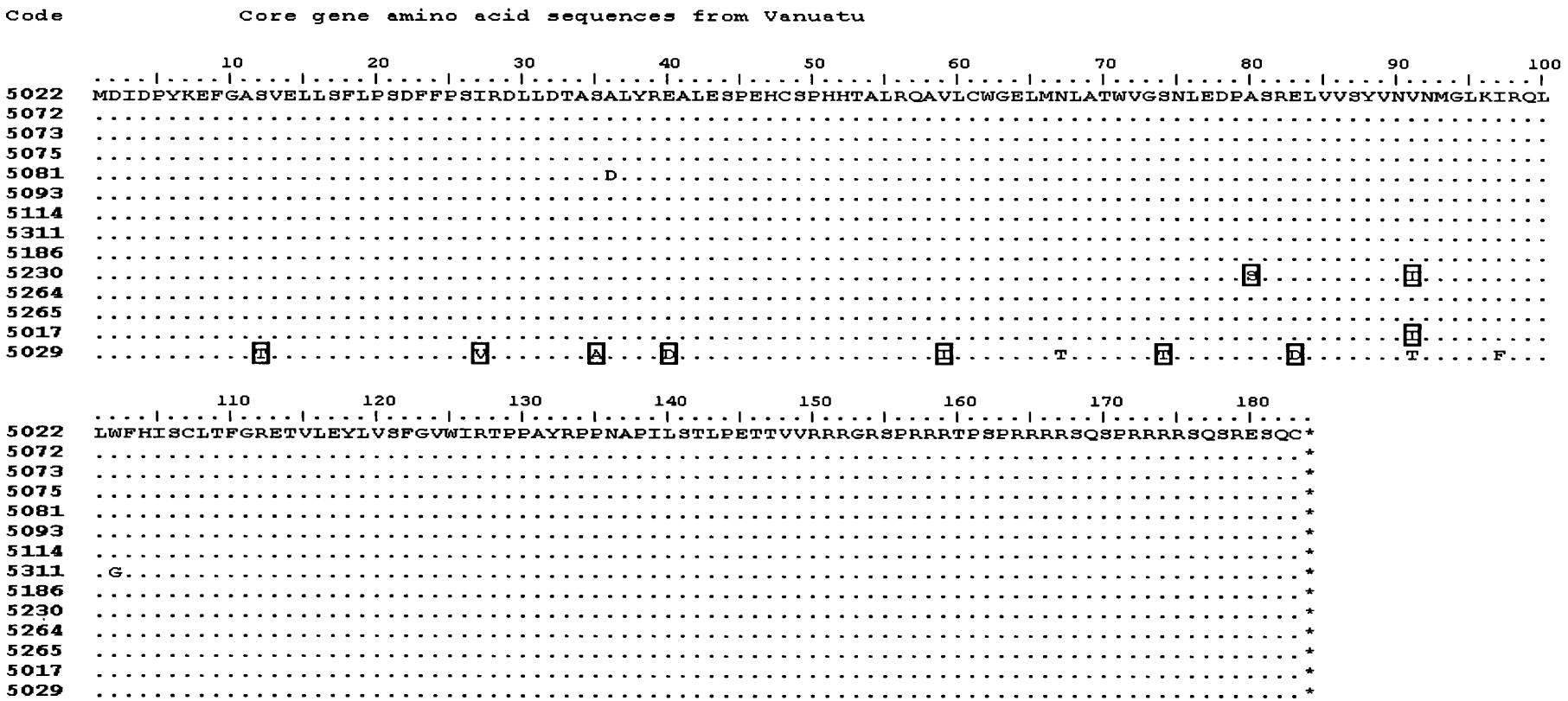


Table A.I.10) Alignment of complete nucleotide sequences from Vanuatu.

```
Code                 Core gene nucleotide sequences from Vanuatu

             10        20        30        40        50        60        70        80        90       100
     ....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|
5022 ATGGACATTGACCCTTATAAAGAATTTGGAGCTTCTGTGGAGTTACTCTCTTTTTTGCCTTCTGATTTCTTTCCGTCCATTCGAGATCTCCTCGACACCG
5072 .............................................................................T..A...................
5073 ....................................................................................................
5075 ....................................................................................................
5081 ....................................................................................................
5093 ...................................C...............................................G................
5114 .............................................................................T......................
5311 .............................C......................................................................
5186 ................................................................................C...................
5230 .............................................................................T......................
5264 ...................................................................................G................
5265 ...................................................................................G................
5017 ..........................................................................A..T......................
5029 .................................A................G..............C........T...G.A........T..A..T....

            110       120       130       140       150       160       170       180       190       200
     ....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|
5022 CCTCAGCTCTGTATCGGGAGGCCTTAGAGTCTCCGGAGCATTGTTCACCTCACCATACAGCACTCAGGCAAGCTGTTCTGTGTTGGGGTGAGTTAATGAA
5072 .......C......A....A.................A..............................................................
5073 ....................................................................................................
5075 ....................................................................................................
5081 ......A.............................................................................................
5093 ....................................................................................................
5114 ....................................................................................................
5311 ....................................................................................................
5186 ....................................................................................................
5230 ....T................................A..............................................................
5264 ....................................................................................................
5265 ....................................................................................................
5017 ........T............................A..............................................................
5029 ..G.G..............T..............T.......................T..............AA....T..C.....A..AC......C
```

Appendix

Code Core gene nucleotide sequences from Vanuatu

```
          210       220       230       240       250       260       270       280       290       300
     ....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|
5022 TCTGGCTACCTGGGTGGGAAGTAATTTGGAAGATCCAGCATCCAGAGAATTAGTAGTCAGTTATGTCAATGTTAATATGGGCCTAAAAATCCGACAACTA
5072 ...A.............T.............C..C........G.....................................................A.....T..
5073 ....................................................................................................
5075 ....................................T.............T.............C.....................A........
5081 ....................................T.............T.............C.....................A........
5093 ....................................................C.............T.......TA........
5114 ....................................................................................A.......C
5311 ....................................T.............T.............C.....................A........
5186 ....................................................................................A........
5230 ..............................C...T.......G........T.............CA....................A........
5264 ....................................................C.....................A........
5265 ....................................................C.....................A........
5017 ..............................C...........G........T.............CA....................A........
5029 ...A.............T.C......A.............T..G..CC..................CAC..............GT..A.....T..

          310       320       330       340       350       360       370       380       390       400
     ....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|
5022 TTGTGGTTTCACATTTCCTGTCTTACTTTTGGAAGAGAAACTGTCCTTGAGTACTTGGTGTCTTTCGGAGTGTGGATTCGCACTCCTCCCGCTTACAGAC
5072 ....................................T........T...........T.....................A........
5073 ....................................................................................................
5075 ....................................T...................................................
5081 ....................................T...................................................G.
5093 ....................................T...................................................
5114 ....................................T...................................................
5311 ...G................................T...................................................
5186 ....................................................................................................
5230 C...................................T........T...........T...................................
5264 ....................................T...................................................
5265 ....................................T...................................................
5017 ....................................T........T...........T...................................
5029 ................T.....C.............G..T..A.....T...........T..................C..A.....T....
```

Appendix

Code Core gene nucleotide sequences from Vanuatu

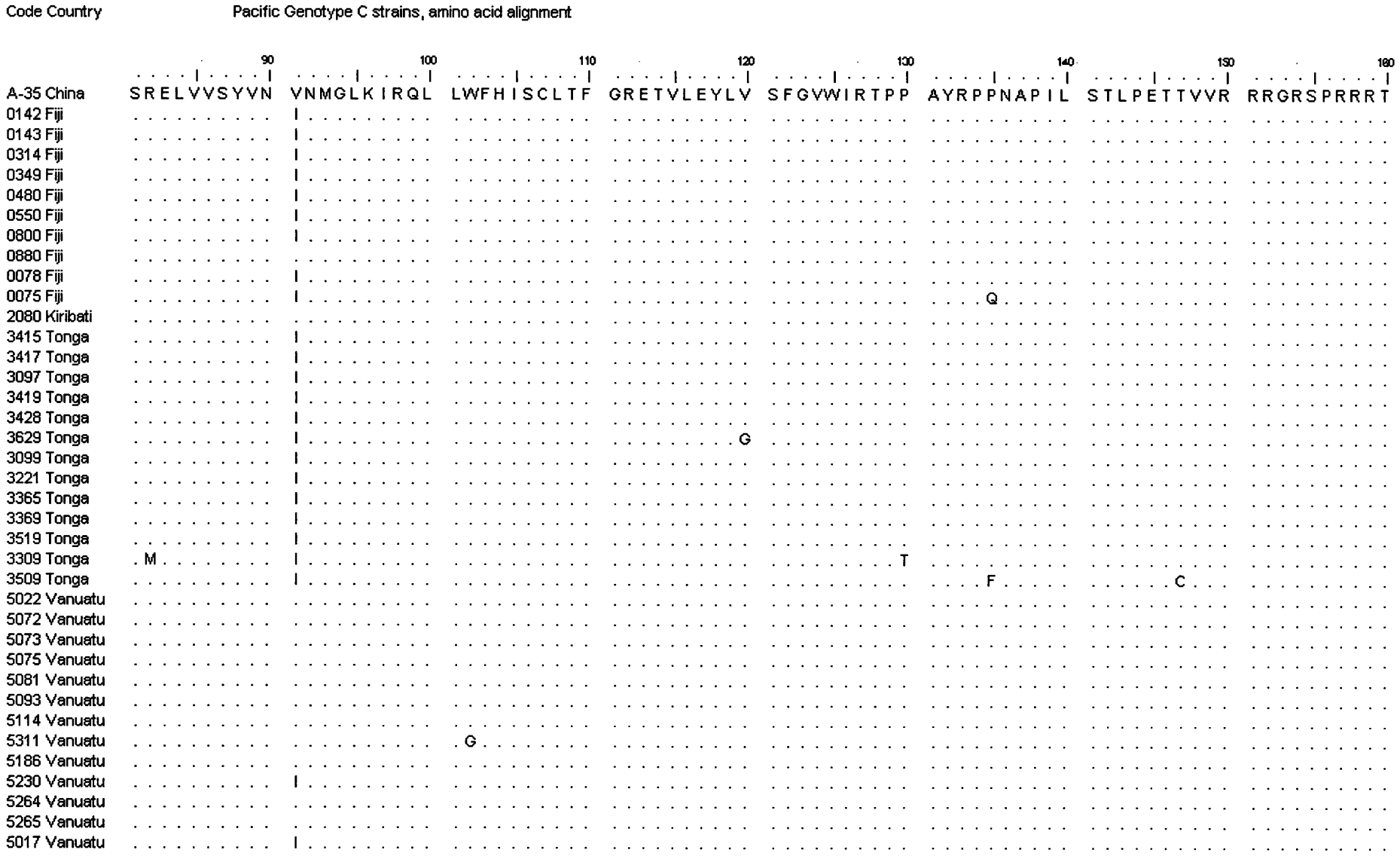
```
          410       420       430       440       450       460       470       480       490       500
     ....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|
5022 CACCAAATGCCCCTATCTTATCAACACTTCCGGAAACTACTGTTGTTAGACGACGAGGCAGGTCCCCTAGAAGAAGAACTCCCTCGCCTCGCAGACGAAG
5072 ..................................................................................A.................
5073 ....................................................................................................
5075 ....C.....................................................A.....................................
5081 ..........................................................A..A.....................................
5093 .....................................................A.....................................
5114 .....................................................A.....................................
5311 ................T....................................A.....................................A....
5186 ....................................................................................................
5230 ....................................................................................................
5264 .....................................................A.....................................A....
5265 .....................................................A.....................................
5017 ....................................................................................................
5029 ................C.................G.....................................................

          510       520       530       540       550
     ....|....|....|....|....|....|....|....|....|....|..
5022 ATCTCAATCGCCGCGTCGCAGAAGATCTCAATCTCGGGAATCTCAATGTTAG
5072 G...................................................
5073 ...................................................
5075 ...................................................
5081 G........A..........................................
5093 ...................................................
5114 ...................................................
5311 ...................................................
5186 ...................................................
5230 .........A.....................................
5264 ...................................................
5265 ...................................................
5017 G........A..........................................
5029 ...................................................
```

Table A.I.11) Alignment of complete HBcAg amino acid sequences of adrq- strains (genotype C) from the Pacific islands.

Code	Country	Pacific Genotype C strains, amino acid alignment
		10 20 30 40 50 60 70 80
A-35	China	MDIDPYKEFG ASVELLSFLP SDFFPSIRDL LDTASALYRE ALESPEHCSP HHTALRQAIL CWGELMNLAT WVGSNLEDPA
0142	Fiji	. . . . . V . . . . .
0143	Fiji	. . . . . V . . . . .
0314	Fiji	. . . . . V . . . . .
0349	Fiji	. . . . . V . . . . .
0480	Fiji	. . . . . V . . . . .
0550	Fiji	. . . . . V . . . . .
0800	Fiji	. . . . . V . . . . .
0880	Fiji	. . . . . V . . . . .
0078	Fiji	. . . . . V . . . . .
0075	Fiji	. . . . . P . A . . . . . C . . . . . V . . . . .
2080	Kiribati	. . . . . V . . . . .
3415	Tonga	. . . . . V . . . . .
3417	Tonga	. . . . . V . . . . .
3097	Tonga	. . . . . V . . . . .
3419	Tonga	. . . . . V . . . . .
3428	Tonga	. . . . . V . . . . .
3629	Tonga	. . . . . V . . . . .
3099	Tonga	. . . . . V . . . . .
3221	Tonga	. . . . . V . . . . .
3365	Tonga	. . . . . V . . . . .
3369	Tonga	. . . . . V . . . . .
3519	Tonga	. . . . . V . . . . .
3309	Tonga	. . . . . V . W . . . . . R S . D . . T
3509	Tonga	. . . . . V . . . . . D . . . . .
5022	Vanuatu	. . . . . V . . . . .
5072	Vanuatu	. . . . . V . . . . .
5073	Vanuatu	. . . . . V . . . . .
5075	Vanuatu	. . . . . V . . . . .
5081	Vanuatu	. . . . . D . . . . . V . . . . .
5093	Vanuatu	. . . . . V . . . . .
5114	Vanuatu	. . . . . V . . . . .
5311	Vanuatu	. . . . . V . . . . .
5186	Vanuatu	. . . . . V . . . . .
5230	Vanuatu	. . . . . V . . . . . S
5264	Vanuatu	. . . . . V . . . . .
5265	Vanuatu	. . . . . V . . . . .
5017	Vanuatu	. . . . . V . . . . .

Appendix


Code Country
Pacific Genotype C strains, amino acid alignment
90
100
110
120
130
140
150
160
A-35 China SRELVVSYVN VNMGLKIRQL LWFHISCLTF GRETVLEYLV SFGVWIRTPP AYRPPNAPIL STLPETTVVR RRGRSPRRRT
0142 Fiji
0143 Fiji
0314 Fiji
0349 Fiji
0480 Fiji
0550 Fiji
0800 Fiji
0880 Fiji
0078 Fiji
0075 Fiji
2080 Kiribati
3415 Tonga
3417 Tonga
3097 Tonga
3419 Tonga
3428 Tonga
3629 Tonga
3099 Tonga
3221 Tonga
3365 Tonga
3369 Tonga
3519 Tonga
3309 Tonga
3509 Tonga
5022 Vanuatu
5072 Vanuatu
5073 Vanuatu
5075 Vanuatu
5081 Vanuatu
5093 Vanuatu
5114 Vanuatu
5311 Vanuatu
5186 Vanuatu
5230 Vanuatu
5264 Vanuatu
5265 Vanuatu
5017 Vanuatu

Appendix

Code Country — Pacific Genotype C strains, amino acid alignment

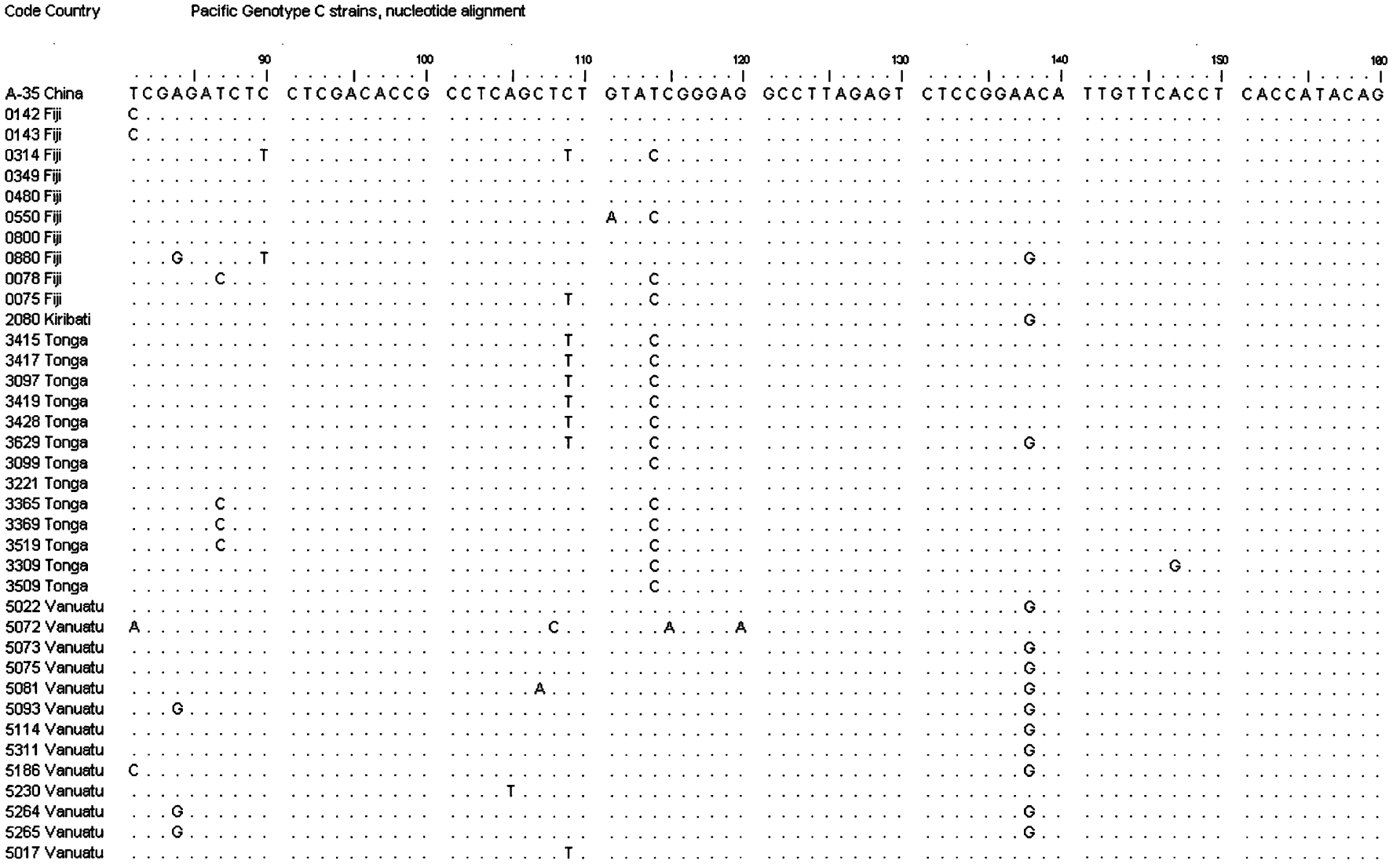
	170									180								
A-35 China	P	S	P	R	R	R	R	S	Q	S	P	R	R	R	R	S	Q	S	R	E	S	Q	C	*
0142 Fiji	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	*
0143 Fiji	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	*
0314 Fiji	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	*
0349 Fiji	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	*
0480 Fiji	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	*
0550 Fiji	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	*
0800 Fiji	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	*
0880 Fiji	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	*
0078 Fiji	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	*
0075 Fiji	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	*
2080 Kiribati	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	*
3415 Tonga	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	*
3417 Tonga	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	*
3097 Tonga	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	*
3419 Tonga	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	*
3428 Tonga	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	*
3629 Tonga	.	.	.	.	.	.	.	.	.	.	P	.	.	.	.	.	.	.	.	.	.	.	.	*
3099 Tonga	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	*
3221 Tonga	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	*
3365 Tonga	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	*
3369 Tonga	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	*
3519 Tonga	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	*
3309 Tonga	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	*
3509 Tonga	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	*
5022 Vanuatu	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	*
5072 Vanuatu	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	*
5073 Vanuatu	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	*
5075 Vanuatu	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	*
5081 Vanuatu	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	*
5093 Vanuatu	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	*
5114 Vanuatu	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	*
5311 Vanuatu	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	*
5186 Vanuatu	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	*
5230 Vanuatu	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	*
5264 Vanuatu	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	*
5265 Vanuatu	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	*
5017 Vanuatu	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	*

Table A.I.12) Alignment of complete nucleotide sequences of adrq- strains (genotype C) from the Pacific islands.

Code Country — Pacific Genotype C strains, nucleotide alignment

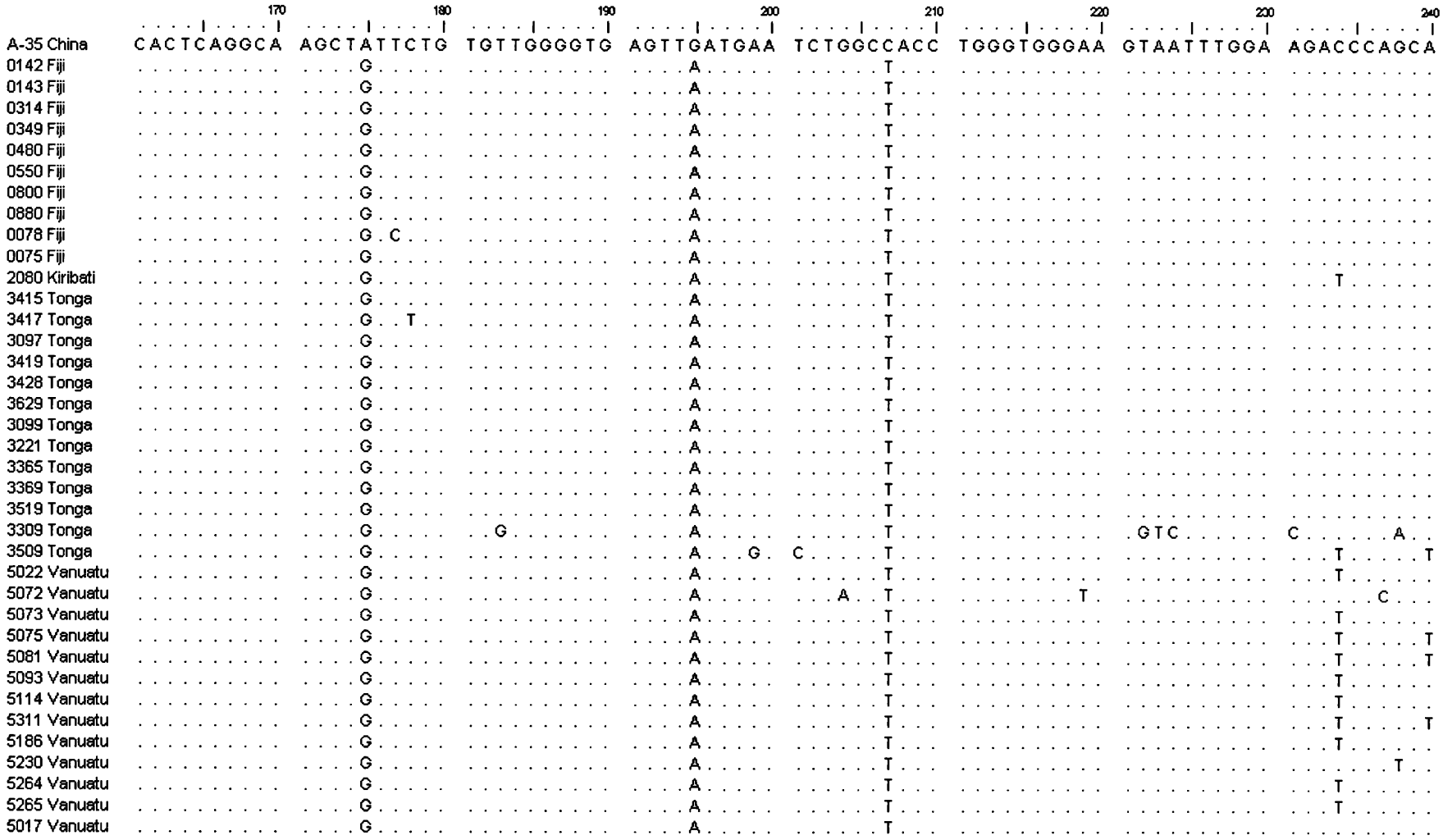
	10	20	30	40	50	60	70	80
A-35 China	ATGGACATTG	ACCCGTATAA	AGAATTTGGA	GCTTCTGTGG	AGTTACTCTC	TTTTTTGCCT	TCTGACTTCT	TTCCTTCTAT
0142 Fiji		T					T	C G
0143 Fiji		T					T	C G
0314 Fiji		T		C			T	G
0349 Fiji		T					T	C G
0480 Fiji		T					T	G
0550 Fiji		A					T	G
0800 Fiji		T					T	G
0880 Fiji		T					T	G C
0078 Fiji		T					T	G
0075 Fiji		T		C C	C		T	G
2080 Kiribati		T					T	G C
3415 Tonga		T		C			T	G
3417 Tonga		T		C C			T	G
3097 Tonga		T		C			T	G
3419 Tonga		T		C			T	G
3428 Tonga		T		C			T	G
3629 Tonga		T		C	A		T	G
3099 Tonga		T					T	G C
3221 Tonga		T					T	G
3365 Tonga		T					T	G
3369 Tonga		T					T	G
3519 Tonga		T					T	G
3309 Tonga		T					T	G
3509 Tonga		T					T	C G
5022 Vanuatu		T					T	G C
5072 Vanuatu		T					T	G
5073 Vanuatu		T					T	G C
5075 Vanuatu		T					T	G C
5081 Vanuatu		T					T	G C
5093 Vanuatu		T		C			T	G C
5114 Vanuatu		T					T	G
5311 Vanuatu		T	C				T	G C
5186 Vanuatu		T					T	G C
5230 Vanuatu		T					T	G
5264 Vanuatu		T					T	G C
5265 Vanuatu		T					T	G C
5017 Vanuatu		T					T	A

Appendix


Code Country
Pacific Genotype C strains, nucleotide alignment
90
100
110
120
130
140
150
160
A-35 China TCGAGATCTC CTCGACACCG CCTCAGCTCT GTATCGGGAG GCCTTAGAGT CTCCGGAACA TTGTTCACCT CACCATACAG
0142 Fiji
0143 Fiji
0314 Fiji
0349 Fiji
0480 Fiji
0550 Fiji
0800 Fiji
0880 Fiji
0078 Fiji
0075 Fiji
2080 Kiribati
3415 Tonga
3417 Tonga
3097 Tonga
3419 Tonga
3428 Tonga
3629 Tonga
3099 Tonga
3221 Tonga
3365 Tonga
3369 Tonga
3519 Tonga
3309 Tonga
3509 Tonga
5022 Vanuatu
5072 Vanuatu
5073 Vanuatu
5075 Vanuatu
5081 Vanuatu
5093 Vanuatu
5114 Vanuatu
5311 Vanuatu
5186 Vanuatu
5230 Vanuatu
5264 Vanuatu
5265 Vanuatu
5017 Vanuatu

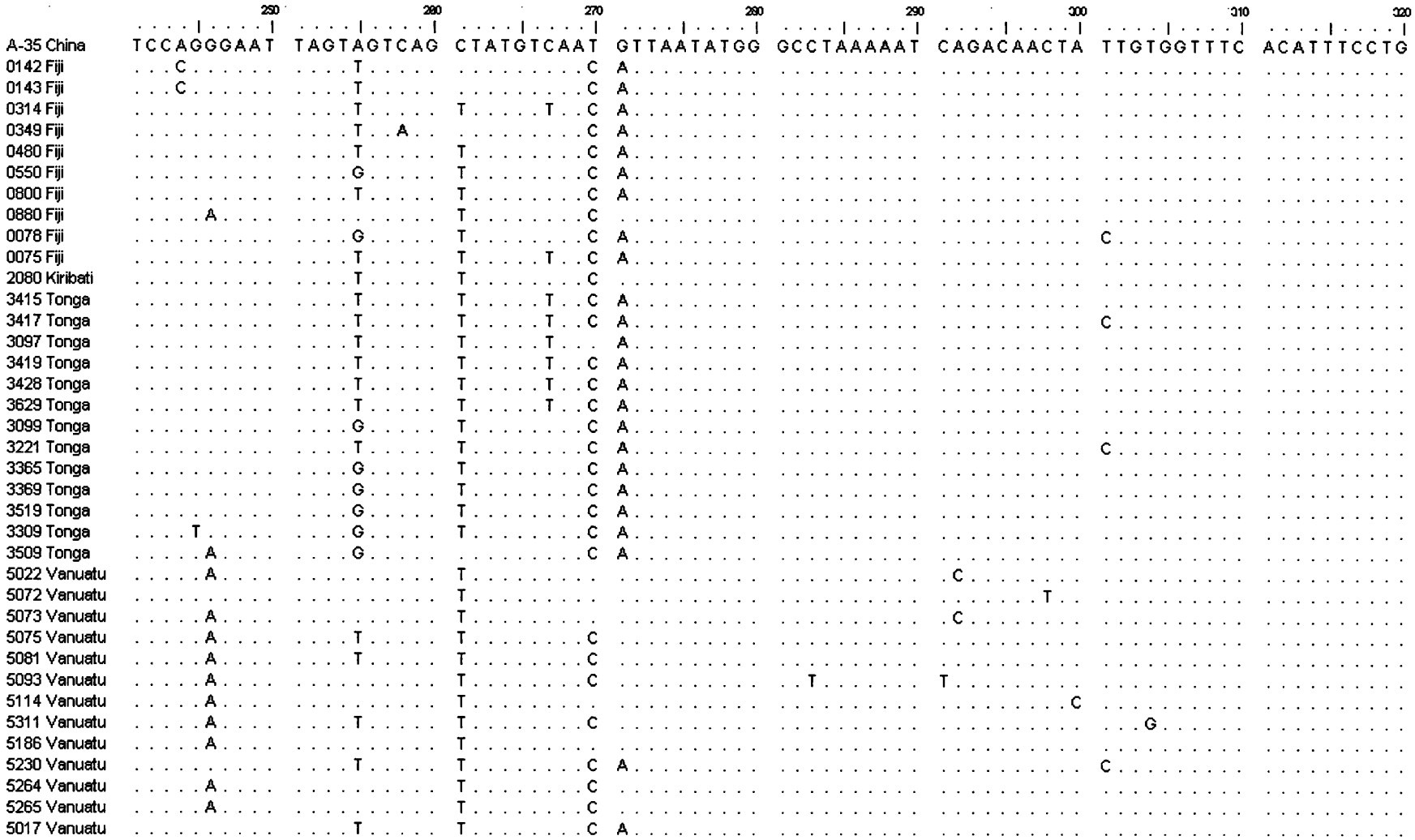
Appendix

Code Country Pacific Genotype C strains, nucleotide alignment



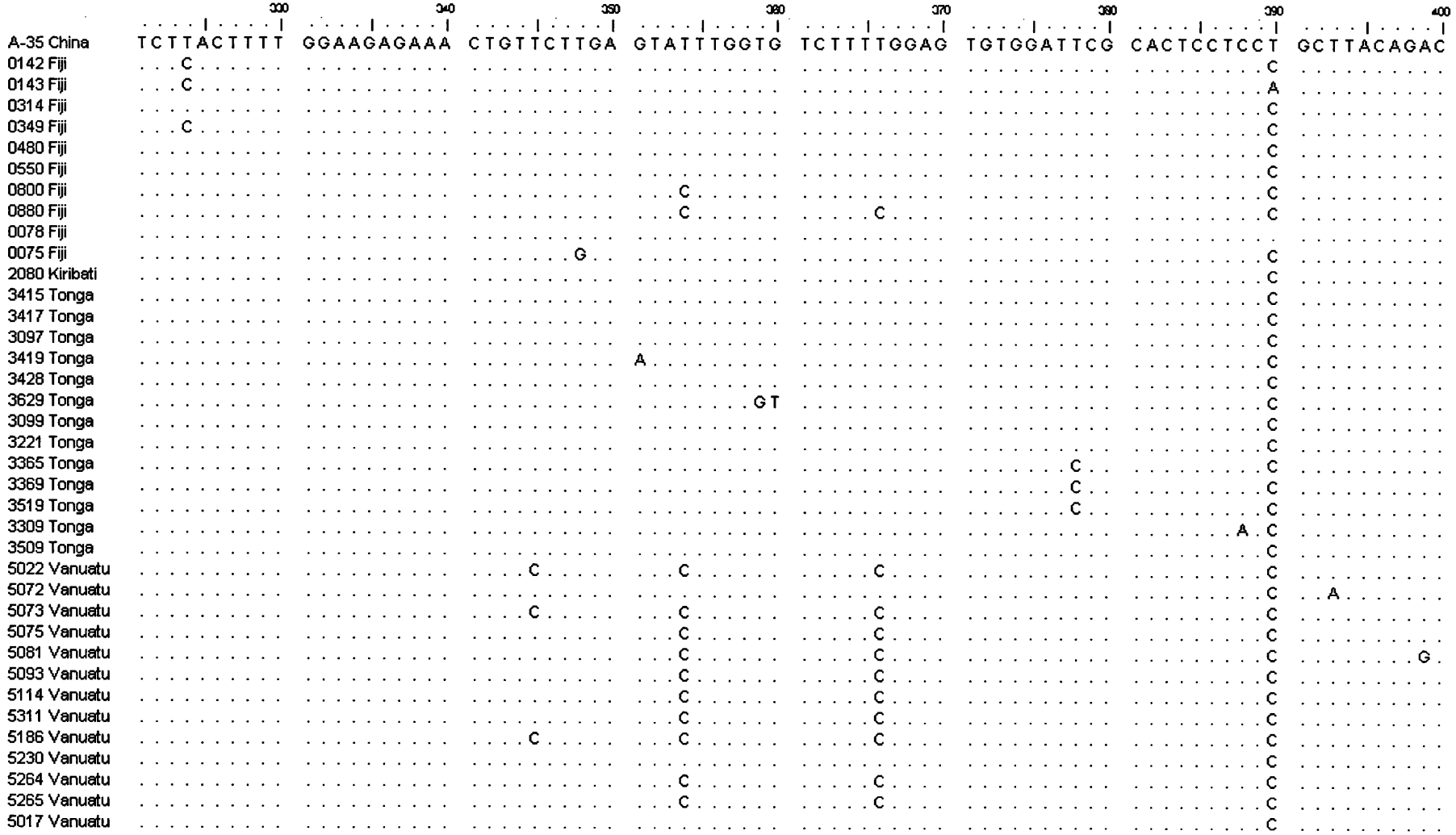
Appendix

Code Country Pacific Genotype C strains, nucleotide alignment



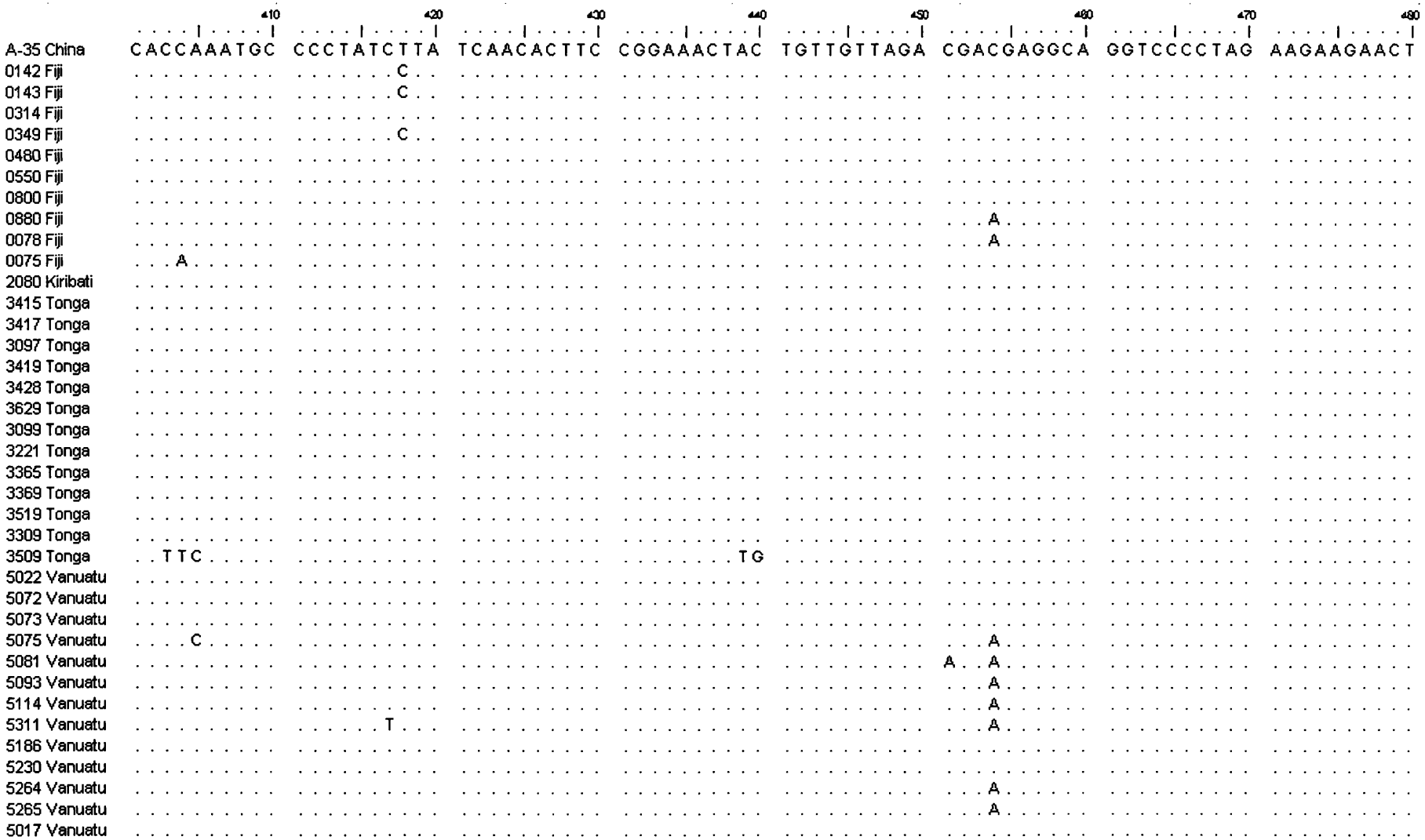
Appendix

Code Country Pacific Genotype C strains, nucleotide alignment



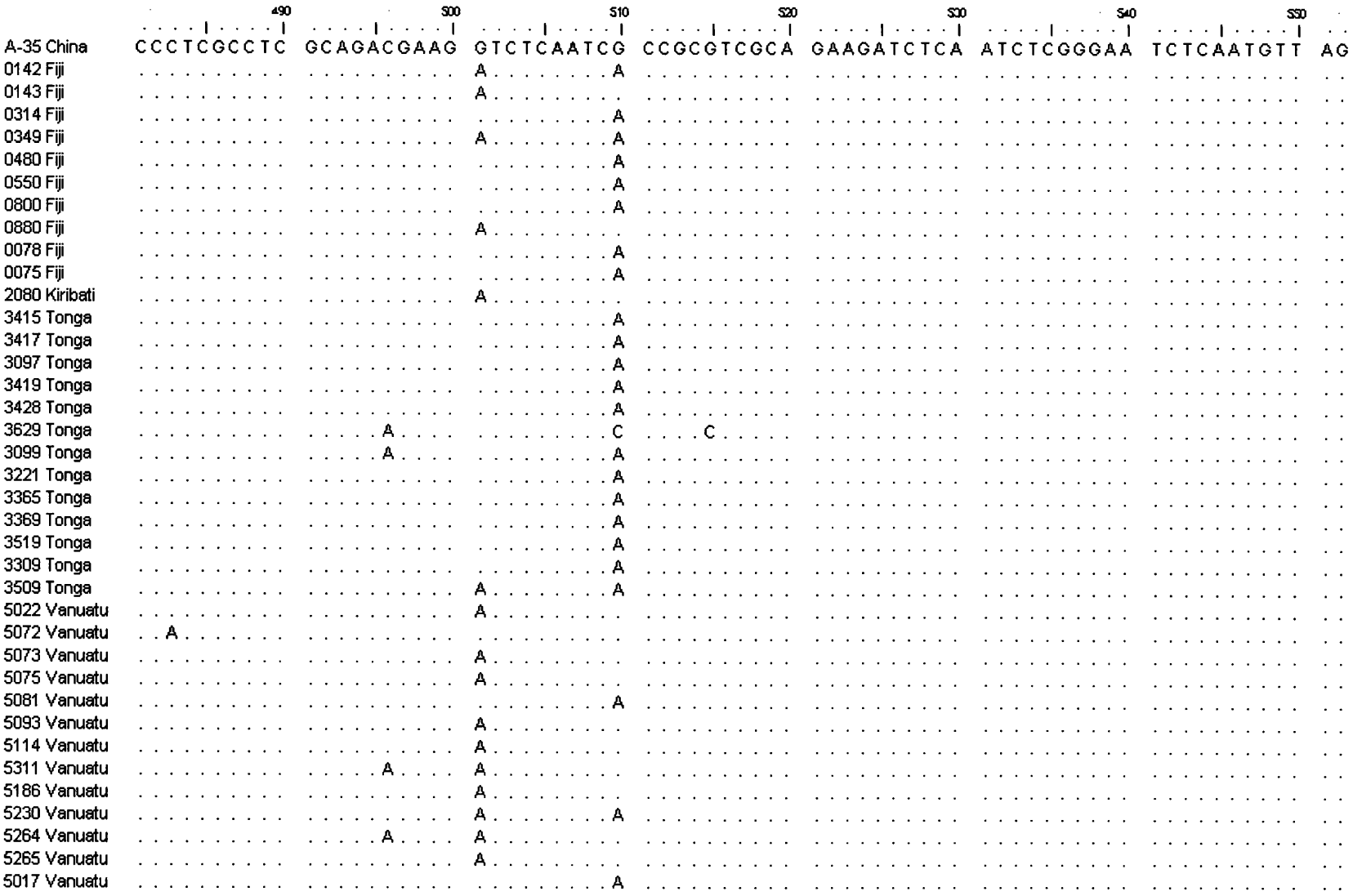
Appendix

Code Country Pacific Genotype C strains, nucleotide alignment



Appendix

Code Country Pacific Genotype C strains, nucleotide alignment



Appendix

Note: Strains belonging to Vanuatu showed a divergent nucleotide alignment (b) with a characteristic V91 at the aa levels. Amino acids and nucleotides are numbered from the beginning of the HBcAg.

Table A.I.13) Alignment of complete HBcAg amino acid sequences of ayw (genotype D) strains from Pacific islands.

Appendix

Code Country Pacific genotype D strains, amino acid alignment

```
                    10        20        30        40        50        60        70        80        90       100
                    |....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|
0123 India  MDIDPYKEFG ATVELLSFLP SDFFPSVRDL LDTASALYRE ALESPEHCSP HHTALRQAIL CWGELMTLAT WVGGNLEDPI SRDLVVSYVN TNMGLKFRQL
2006 Kirib .......... .......... .......... ....A....D .......... .......... .......... ...T.....A .......... ..........
2007 Kirib .......... .......... .......... ....A....D .......... .......... .......... ...T.....A .......... ..........
2019 Kirib .......... .......... .......... ....A....D .......... .......... .......... ...T.....A .......... ..........
2119 Kirib .......... .......... .......... ....A....D .......... .......... .......... ...T.....A .......... ..........
2127 Kirib .......... .......... .......... ....A....D .......... .......... .......... ...T.....A .......... ..........
2317 Kirib .......... .......... .......... ....A....D .......... .......... .......... ...T....QA .......... ..........
2483 Kirib .......... .......... .......... ....A....D .........S .......... .......... ...T.....A .......... ..........
2109 Kirib .......... .......... .......... ....A....D .......... .......... .......... ...T.....A .......... ..........
2110 Kirib .......... .......... .......... ....A....D .......... .......... .......... ...T.....A .......... ..........
2117 Kirib .......... .......... .......... ....A....D .......... .......... .......... ...T.....A .......... ..........
2084 Kirib .......... .......... .......... ....A....D .......... .......... .......... ...T.....A .......... ..........
2039 Kirib .......... .......... .......... ....A....D .......... .......... .......... ...T.....A .......... ..........
2143 Kirib ..........  ..GSV..... .......... ....A....D .........A .......... .......... ...T..XX.A .......... ..........
0079 Fiji  .......... .......... .......... ....A....D .......... .......... .......... ...T.....A .......... ..........
0092 Fiji  .......... .......... .......... ....A....D .......... .......... .......... ...T.....A .......... ..........
0470 Fiji  .......... .......... .......... ....A....D .......... .......... .......... ...T.....A .......... ..........
0456 Fiji  .......... .......... .......... ....A....D .......... .......... .......... ...T.....A .......... ..........
0026 Fiji  .......... .......... .......... ....A....D .......... .......... .......... ...T.....A .......... ..........
3343 Tonga .......... .......... .......... ....A....D .......... .......... .......... ...T.....A .......... ..........
3414 Tonga .......... .......... .......... ....A....D .......... .......... .......... ...T.....A .......... ..........
5029 Vanua .......... .......... .......... ....A....D .......... .......... .......... ...T.....A .......... ..........

                   110       120       130       140       150       160       170       180
                    |....|....|....|....|....|....|....|....|....|....|....|....|....|....|....
0123 India  LWFHISCLTF GRETVLEYLV SFGVWIRTPP AYRPPNAPIL STLPETTVVR RRGRSPRRRT PSPRRRRSQS PRRRRSQSRE SQC*
2006 Kirib .......... .....L.... .......... .......... .......... .......... .......... .......... ...*
2007 Kirib .......... .....L.... .......... .......... .......... .......... .......... .......... ...*
2019 Kirib .......... .....L.... .......... .......... .......... .......... .......... .......... ...*
2119 Kirib .......... .....L.... .......... .......... .......... .......... .......... .......... ...*
2127 Kirib .......... .....L.... .......... .......... .......... .......... .......... .......... ...*
2317 Kirib .......... .....L.... .........Q .......... ......C.T. .......... .......... .......... ...*
2483 Kirib .......... .....L.... .......... .......... .......... T......... .......... .......... ...*
2109 Kirib .......... .....L.... .......... .......... .......... .......... .......... .......... ...*
2110 Kirib .......... .....L.... .......... .......... .......... .......... .......... .......... ...*
2117 Kirib .......... .....L.... .......... .......... .......... .......... .......... .......... ...*
2084 Kirib .......... .....L.... .......... .......... .......... .......... .......... .......... ...*
2039 Kirib .......... .....L.... .......... .......... .......... .......... .......... .......... ...*
2143 Kirib .......... .....[.]. .......... .......... .......... .......... .......... .......... ...*
0079 Fiji  .......... .....L.... .......... .......... .......... .......... .......... .......... ...*
0092 Fiji  .......... .....L.... .......... .......... .......... .......... .......... .......... ...*
0470 Fiji  .......... .....L.... .......... .......... .......... .......... .......... .......... ...*
0456 Fiji  .......... .....L.... .......... .......... .......... .......... .......... .......... ...*
0026 Fiji  .......... .....L.... .......... .......... .......... .......... .......... .......... ...*
3343 Tonga .......... .....L.... .......... .......... ....A..... .......... .......... .......... ...*
3414 Tonga .......... .....L.... .......... .......... .......... .......... .......... .......... ...*
5029 Vanua .......... .....L.... .......... .......... .......... .......... .......... .......... ...*
```

Appendix

Table A.I.14) Alignment of complete nucleotide sequences of ayw (genotype D) strains from Pacific islands.

Code Country — Pacific genotype D strains, nucleotide alignment

```
                    10        20        30        40        50        60        70        80        90       100
                    |....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|
0123 India          ATGGACATTG ATCCTTATAA AGAATTTGGA GCTACTGTGG AGTTACTCTC GTTTCTTCCT TCTGACTTCT TTCCATCAGT ACGAGATCTT CTAGATACCG
2006 Kirib          .......... .C........ .......... .......... .......... ....T.G... .......... ....T..C.. .......... ..........
2007 Kirib          .......... .C........ .......... .......... .......... ....T.G... .......... ....T..C.. .......... ..........
2019 Kirib          .......... .C........ .......... .......... .......... ....T.G... .......... ....T..C.. .......... ..........
2119 Kirib          .......... .C........ .......... .......... .......... ....T.G... .......... ....T..C.. .......... ..........
2127 Kirib          .......... .C........ .......... .......... .......... ....T.G... .......... ....T..C.. .......... ..........
2317 Kirib          .......... .C..G..... .......... .......... .......... ....T.G... .......... ....T..C.. .......... ..........
2483 Kirib          .......... .C........ .......... .......... .......... ....T.G... .........T. ....T..T.. .......... .....C....
2109 Kirib          .......... .C........ .......... .......... .......... ....T.G... .......... ....T..C.. .......... ..........
2110 Kirib          .......... .C........ .......... .......... .......... ....T.G... .......... ....T..C.. .......... ..........
2117 Kirib          .......... .C........ .......... .......... .......... ....T.G... .......... ....T..C.. .......... ..........
2084 Kirib          .......... .C........ .......... .......... .......... ....T.G... .......... ....T..C.. .......... ..........
2039 Kirib          .......... .C........ .......... .......... .......... ....T.G... .......... ....T..C.. .......... ..........
2143 Kirib          .......... .C........ .......... .......G.A GTG....... ....T.G... .......... ....T..C.. .......... ..........
0079 Fiji           .......... .C........ .......... .......... .......... ....T.G... .......... ....T..C.. .......... ..........
0092 Fiji           .......... .C........ .......... .......... .......... ....T.G... .......... ....T..C.. .......... ..........
0470 Fiji           .......... .C........ .......... .......... .......... ....T.G... .......... ....T..C.. .......... ..........
0456 Fiji           .......... .C........ .......... .....C.... .......... ....T.G... .......... ....T..C.. .......... ..........
0026 Fiji           .......... .C........ .......... .......... .......... ....T.G... .......... ....T..C.. .......... ..........
3343 Tonga          .......... .C........ .......... .......... .......... ....T.G... .......... ....T..C.. .......... ..........
3414 Tonga          .......... .C........ .......... .......... .......... ....T.G... .......... ....T..C.. .......... ..........
5029 Vanua          .......... .C........ .......... .......... .......... ....T.G... .......... ....T..C.. .......... ..........

                    110       120       130       140       150       160       170       180       190       200
                    |....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|
0123 India          CCTCAGCTCT ATATCGGGAA GCCTTAGAGT CGCCTGAACA TTGTTCACCT CACCATACTG CACTCAGGCA AGCAATTCTT TGCTGGGGGG AACTAATGAC
2006 Kirib          ..G.G..... G........T .......... .T.....G.. .......... .......... .......... .......... ........A. ..........
2007 Kirib          ..G.G..... G........T .......... .T.....G.. .......... .......... .......... .......... ........A. ..........
2019 Kirib          ..G.G..... G........T .......... .T.....G.. .......... .......... .......... .......... ........A. ..........
2119 Kirib          ..G.G..... G........T .......... .T.....G.. .......... .......... .......... .......... ........A. ..........
2127 Kirib          ..G.G..... G........T .......... .T.....G.. .......... .......... .......... .......... ........A. ..........
2317 Kirib          ..G.G..... G........T .......... .T.....G.. .......... .......... .......... .......... ........A. ..........
2483 Kirib          ..G.G..... G........T .........A. .T.....G.. .......T.. .......... .......... .......... ........A. ..........
2109 Kirib          ..G.G..... G........T .......... .T.....G.. .......... .......... .......... .......... ........A. ..........
2110 Kirib          ..G.G..... G........T .......... .T.....G.. .......... .......... .......... .......... ........A. ..........
2117 Kirib          ..G.G..... G........T .......... .T.....G.. .......... .......... .......... .......... ........A. ..........
2084 Kirib          ..G.G..... G........T .......... .T.....G.. .......... ........C. .......... .......... ........A. ..........
2039 Kirib          ..G.G..... G........T .......... .T.....G.. .......... .......... .......... .......... ........A. ..........
2143 Kirib          ..G.G..... G........T .......... .T.....G.. .......G.. .......... .......... .......... ........A. ..........
0079 Fiji           ..G.G..... G........T .......... .T.....G.. .......... .......... .......... .......... ........A. ..........
0092 Fiji           ..G.G..... G........T .......... .T.....G.. .......... .......... .......... .......... ........A. ..........
0470 Fiji           ..G.G..... G........T ..T..G.... .T.....G.. .......... .......... .......... .......... ........A. ..........
0456 Fiji           ..G.G..... G........T .......... .T.....G.. .......... .......... .......... .......... ........A. ..........
0026 Fiji           ..G.G..... G........T .......... .T.....G.. .......... .......... .......... .......... ........A. ..........
3343 Tonga          ..G.G..... G........T .......... .T.....G.. .......... .......... .......... .......... ........A. ..........
3414 Tonga          ..G.G..... G........T .......... .T.....G.. .......... ..T....... .......... .......... ........A. ..........
5029 Vanua          ..G.G..... G........T .......... .T.....G.. .......... .......... .......... .......... ........A. ..........
```

Appendix

Pacific genotype D strains, nucleotide alignment

Code	Country	210	220	230	240	250	260	270	280	290	300
0123	India	TCTAGCCACC	TGGGTGGGTG	GTAATTTGGA	AGATCCAATA	TCCAGGGACC	TAGTGGTCAG	TTATGTTAAC	ACTAATATGG	GCCTAAAATT	TAGGCAACTA
2006	Kirib	......T...	.........A	C......A..	.......GC.	..T.....	....A.....	......C...	.......G..	C..A...T..	
2007	Kirib	......T...	.........A	C......A..	.......GC.	..T.....	....A.....	......C...	.......G..	C..A...T..	
2019	Kirib	....	.........A	C......A..	.......GC.	..T.....	....A.....	......C...	.......G..	C..A...T..	
2119	Kirib	......T...	.........A	C......A..	.......GC.	..T.....	....A.....	......C...	.......G..	C..A...T..	
2127	Kirib	......T...	.........A	C......A..	.......GC.	..T.....	....A.....	......C...	.......G..	C..A...T..	
2317	Kirib	......T...	.........A	C......A..	.....A.GC.	..T.....	....A.....	......C...	.......G..	C..A...T..	
2483	Kirib	......T...	.........A	C......A..	.......GC.	..T.....	....A.....	......C...	.......G..	C..A...T..	
2109	Kirib	....	.........A	C......A..	.......GC.	..T.....	....A.....	......C...	.......G..	C..A...T..	
2110	Kirib	......T...	.........A	C......A..	.......GC.	..T.....	....A..A..	......C...	.......G..	C..A...T..	
2117	Kirib	.T....T...	.........A	C......A..	.......GC.	..T.....	....A.....	......C...	.......G..	C..A...T..	
2084	Kirib	......T...	.........A	C......A..	.......GC.	..T.....	....A.....	......C...	.......G..	C..A...T..	
2039	Kirib	......T...	.........A	C......A..	.......GC.	..T.....	....A.....	......C...	.......G..	C..A...T..	
2143	Kirib	......T...	.........CA	C......A.-	--.....GC.	..T.....	....A.....	......C...	.......G..	CC.A...T..	
0079	Fiji	......T...	.........A	C......A..	.......GC.	..T.....	....A.....	......C...	.......G..	C..A...T..	
0092	Fiji	......T...	.........A	C......A..	.......GC.	..T.....	....A.....	......C...	.......G..	C..A...T..	
0470	Fiji	......T...	.........A	C......A..	.......GC.	..T.....	....A.....	......C...	.......G..	C..A...T..	
0456	Fiji	......T...	.........A	C......A..	.......GC.	..T.....	....A.....	......C...	.......G..	...A...T..	
0026	Fiji	......T...	.........A	C......A..	.......GC.	..T.....	....A.....	......C...	.......G..	C..A...T..	
3343	Tonga	......T...	.........A	C......A..	.......GC.	..T.....	....A.....	......C...	.......G..	C..A...T..	
3414	Tonga	......T...	.........A	C......A..	.......GC.	..T.....	....A.....	......C...	.......G..	C..A...T..	
5029	Vanua	......T...	.........A	C......A..	.......GC.	..T.....	....A.....	......C...	.......G..	C..A...T..	

Code	Country	310	320	330	340	350	360	370	380	390	400
0123	India	TTGTGGTTTC	ACATTTCTTG	TCTCACTTTT	GGAAGAGAAA	CGGTCATAGA	GTATTTGGTG	TCTTTCGGAG	TGTGGATTCG	CACTCCTCCA	GCTTATAGAC
2006	Kirib	....	....	....	....	....TC....	....	.....T....	....	......C...	....
2007	Kirib	....	....	....	....	....TC....	....	.....T....	....	......C...	....
2019	Kirib	....	....	....	....	....TC....	....	.....T....	....	......C...	....
2119	Kirib	....	....	....	....	....TC....	....	.....T....	....	......C...	....
2127	Kirib	....	....	....	....	....TC....	....	.....T....	....	......C...	....
2317	Kirib	....	....	....	....	....TC....	....	.....T....	....	......C.A.	....
2483	Kirib	....	....	....	....	....TC....	....	.....T....	....	......C...	....
2109	Kirib	....	....	....	....	....TC....	....	.....T....	....	......C...	....
2110	Kirib	....	....	....	....	....TC....	....	.....T....	....	......C...	....
2117	Kirib	....	....	....	....	....TT....	....	.....T....	....	....	....
2084	Kirib	....	....	....	....	....TC....	....	.....T....	....	......C...	....
2039	Kirib	....	....	....	....	....TC....	....	.....T....	....	......C...	....
2143	Kirib	....	....	....	....	....T.....	....	.....T....	....	......C...	....
0079	Fiji	....	....	....	....	....TC....	....	.....T....	....	......C...	....
0092	Fiji	....	....	....	....	....TC....	....	.....T....	....	......C...	....
0470	Fiji	....	....	....	....	....TC....	....	.....T....	....	......C...	....
0456	Fiji	....	....	....	....	....TC....	....	.....T....	....	......C...	....
0026	Fiji	....	....	....	....	....TC....	....	.....T....	....	......C...	....
3343	Tonga	....	....	....	....	....TC....	....	.....T....	....	......C...	....
3414	Tonga	....	....	....	....	....TC....	....	.....T....	....	......C...	....
5029	Vanua	....	....	....	....	....TC....	....	.....T....	....	......C...	....

Appendix

Code Country — Pacific genotype D strains, nucleotide alignment

```
                    410        420        430        440        450        460        470        480        490        500
                    |....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|
0123 India          CACCAAATGC CCCTATCTTA TCAACACTTC CGGAGACTAC TGTTGTTAGA CGACGAGGCA GGTCCCCTAG AAGAAGAACT CCCTCGCCTC GCAGACGAAG
2006 Kirib          ..........C.. .......... .......... .......... .......... .......... .......... .......... ..........
2007 Kirib          ..........C.. .......... .......... .......... .......... .......... .......... .......... ..........
2019 Kirib          ..........C.. .......... .......... .......... .......... .......... .......... .......... ..........
2119 Kirib          ..........C.. .......... .......... .......... .......... .......... .......... .......... ..........
2127 Kirib          ..........C.. .......... .......... .......... .......... .......... .......... .......... ..........
2317 Kirib          ..........C.. .......... .........TG ....A..... .......... .......... .......... .......... ..........
2483 Kirib          ..........C.. .......... .......... .......... .A........ .......... .......... .......... ..........
2109 Kirib          ..........C.. .......... .......... .......... .......... .......... .......... .......... ..........
2110 Kirib          ..........C.. .......... .......... .......... .......... .......... .......... .......... ..........
2117 Kirib          ..........C.. .......... .......... .......... .......... .......... .......... .......... ..........
2084 Kirib          ..........C.. .......... .......... .......... .......... .......... .......... .......... ..........
2039 Kirib          ..........C.. .......... .......... .......... .......... .......... .......... .......... ..........
2143 Kirib          ..........C.. .......... ....A..... .......... .......... .......... .......... .......... ..........
0079 Fiji           ..........C.. .......... .......... .......... .......... .......... .......... .......... ..........
0092 Fiji           ..........C.. .......... .......... .......... .......... .......... .......... .......... ..........
0470 Fiji           ..........C.. .......... .......... .......... .......... .......... .......... .......... ..........
0456 Fiji           ..........C.. .......... ........C.. .......... .......... .......... .......... .......... ..........
0026 Fiji           ..........C.. .......... .......... .......... .......... .......... .......... .......... ..........
3343 Tonga          ..........C.. .......... ...C...... .......... .......... .......... .......... .......... ..........
3414 Tonga          ..........C.. .......... .......... .......... .......... .......... .......... .......... ..........
5029 Vanua          ..........C.. .......... .......... .......... .......... .......... .......... .......... ..........

                    510        520        530        540        550
                    |....|....|....|....|....|....|....|....|....|....|..
0123 India          GTCTCAATCG CCGCGTCGCA GAAGATCTCA ATCTCGGGAA TCTCAATGTT AG
2006 Kirib          A......... .......... .......... .......... .......... ..
2007 Kirib          A......... .......... .......... .......... .......... ..
2019 Kirib          A......... .......... .......... .......... .......... ..
2119 Kirib          A......... .......... .......... .......... .......... ..
2127 Kirib          A......... .......... .......... .......... .......... ..
2317 Kirib          A......... .......... .......... .......... .......... ..
2483 Kirib          A......... .......... .......... .......... .......... ..
2109 Kirib          A......... .......... .......... .......... .......... ..
2110 Kirib          A......... .......... .......... .......... .......... ..
2117 Kirib          A......... .......... .......... .......... .......... ..
2084 Kirib          A......... .......... .......... .......... .......... ..
2039 Kirib          A......... .......... .......... .......... .......... ..
2143 Kirib          .........A .......... .......... .......... .......... ..
0079 Fiji           A......... .......... .......... .......... .......... ..
0092 Fiji           A......... .......... .......... .......... .......... ..
0470 Fiji           A......... .......... .......... .......... .......... ..
0456 Fiji           A......... .......... .......... .......... .......... ..
0026 Fiji           A......... .......... .......... .......... .......... ..
3343 Tonga          A......... .......... .......... .......... .......... ..
3414 Tonga          A......... .......... .......... .......... .......... ..
5029 Vanua          A......... .......... .......... .......... .......... ..
```

A.II Application of core gene sequencing data to an HBV incident

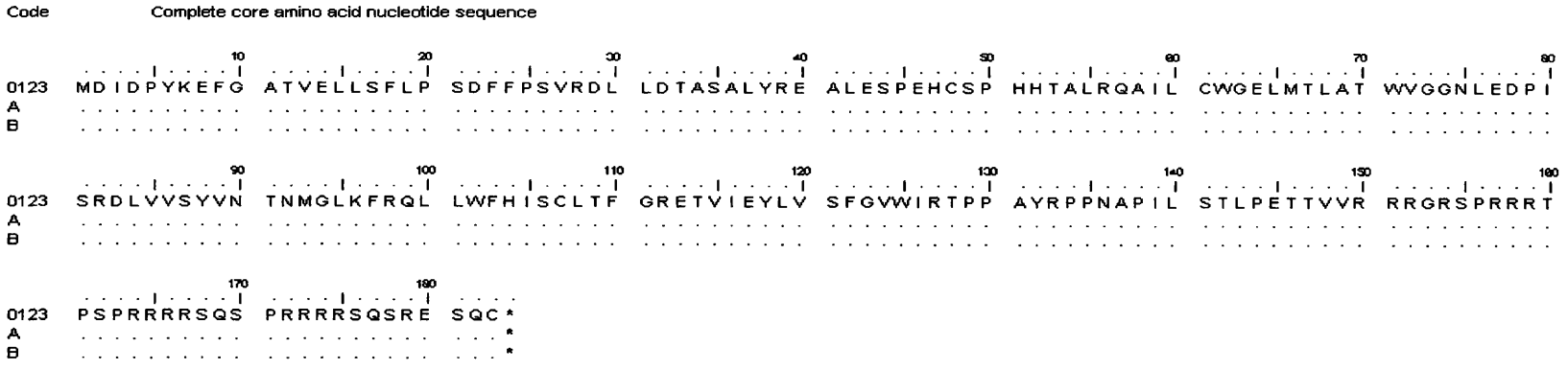
An elderly lady born in the UK, underwent colostomy at the Royal London Hospital in 04.02. She was found to be positive for HBsAg post-operatively (patient B). After an investigation, another HBsAg positive patient was identified. He had had renal transplantation in India in 2001 and underwent colonic surgery at two different London hospitals between 2001 and 2002 (patient A). These two patients had been in different wards; however, both patients were operated on at the same theatre and the same day by the same surgical team. The surgical team were suspected to be a source of HBV in these two patients. However, all member of the team were negative for HBsAg. Stored sera from both patients were sent to the west of Scotland special Virology Centre, Glasgow. DNA from sera was extracted manually (using QIAGEN kit) and a hemi-nested core gene PCR carried out on samples using standard primers (see table 2.1). Sequences were aligned using sequence Navigator program and Bioedit software. One of the Indian core gene sequences (0123) obtained from our previous study (see chapter 3.1) was selected as a reference sequence. The two patient sequences were identical. Compared to the reference Indian sequence (0123) there was 100% homology at amino acid level, and there were only 10 silent variations at the nucleotide level (Tables A.II.1 and A.II.2). These two C gene sequences (Figure A.II.1, bold red sequences, A and B) localised within the Indian cluster in the tree (bold red sequences). Either patient A infected in India by the commonest kind (cluster) of Indian genotype D or in Europe by a form of genotype D that is rare in Europe. Since he was previously in India, the first possibility seems most likely; in which case the Western-derived isolates in the Indian cluster (I55 and 8.25) may also have ultimately come from India (see chapter 3.3).

This HBV outbreak illustrates the value of generating databases of sequences from around the world. Just having the genotype would have been useful but databases of sequences

from this approach allowed us to confirm that the origin of the virus was probably India. However the route of transmission is not yet clear.

Appendix

Table A.II.1. Alignment of complete core gene amino acid sequences of patients A and B. Sample 0123 (from India) used as the reference sequence.



Appendix

Table A.II.2. Alignment of complete core gene nucleotide sequences of patients A and B. Sample 0123 (from India) used as the reference sequence.

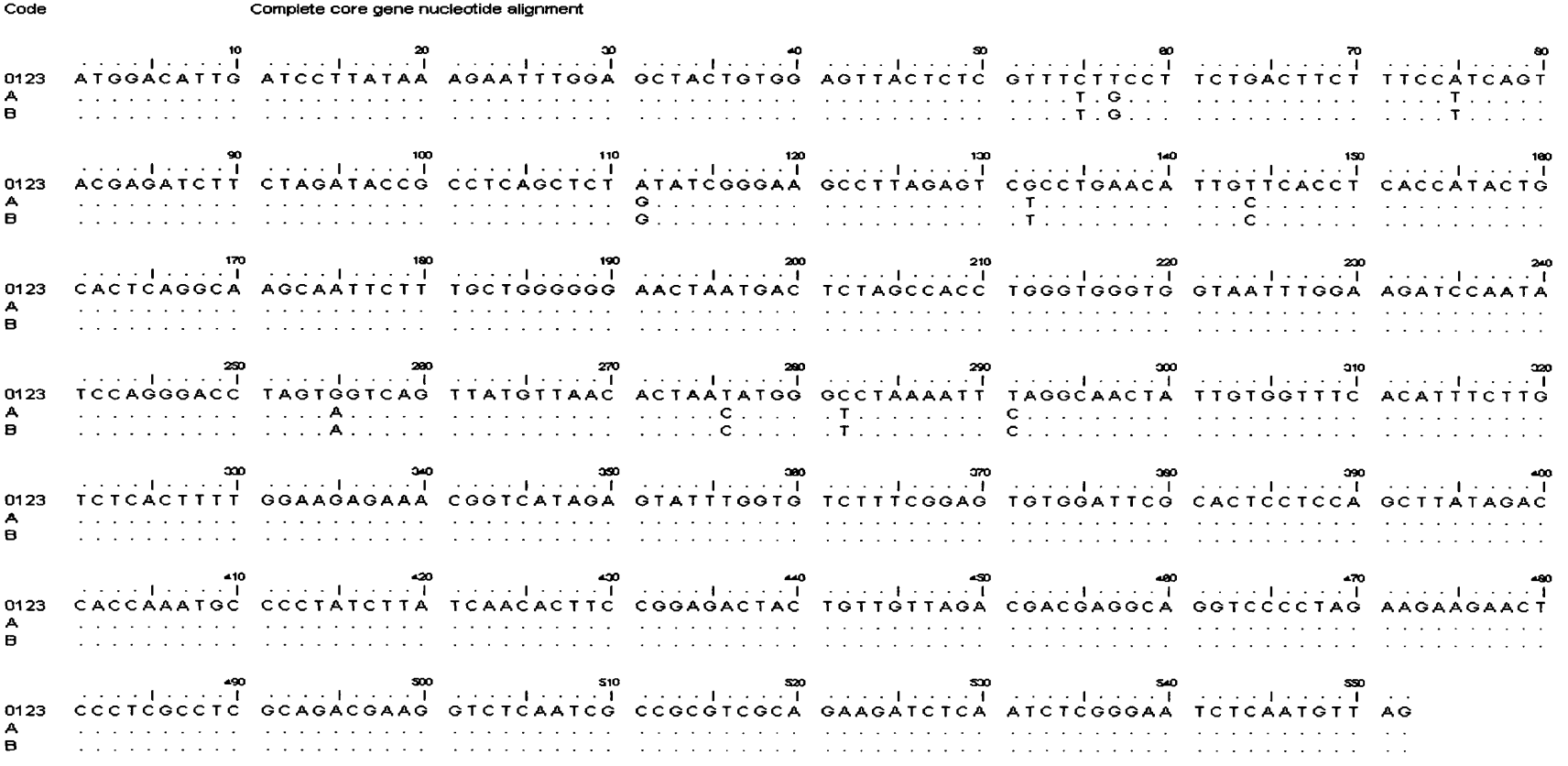
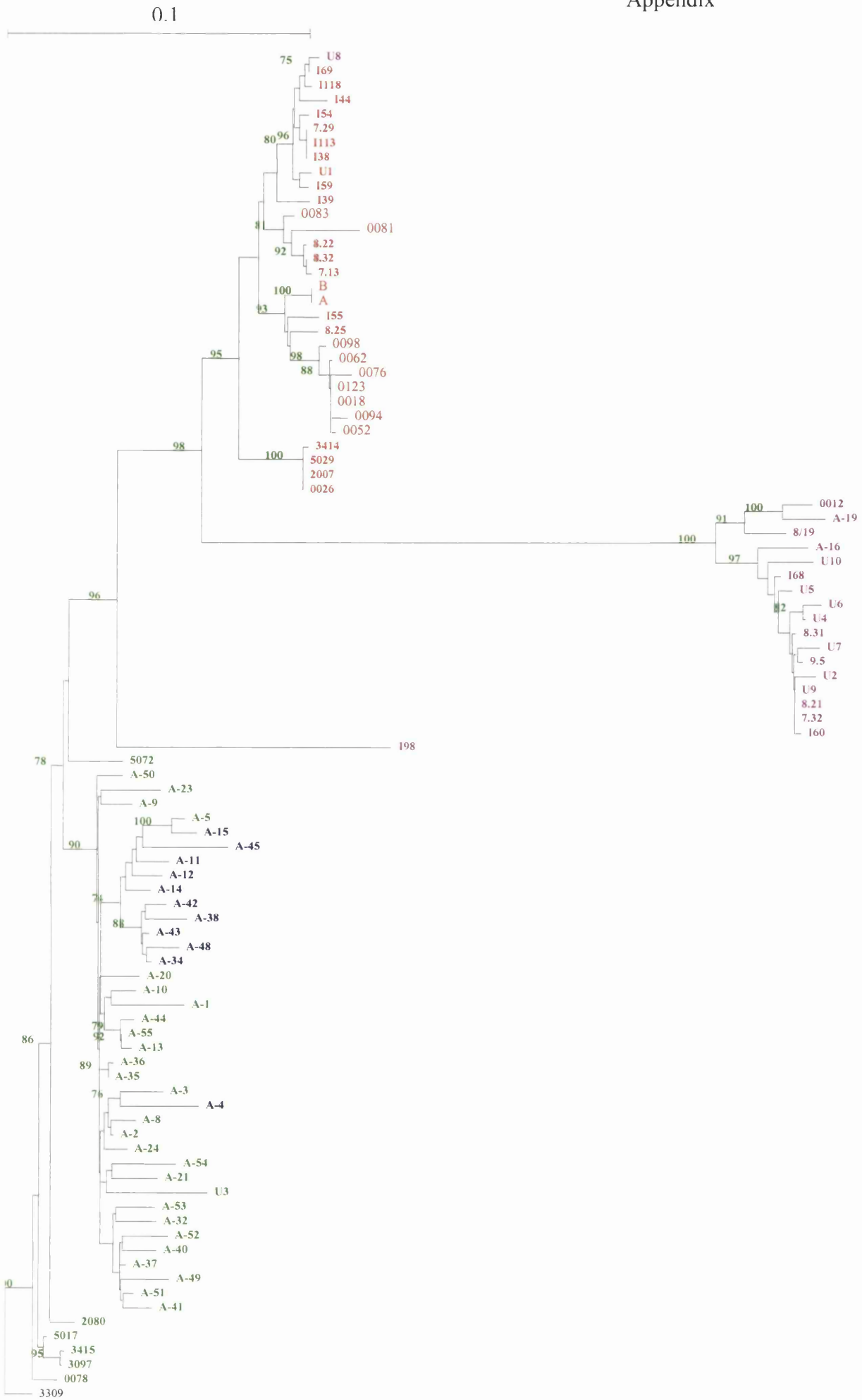


Fig A.II.1. Neighbor joining phylogenetic tree of core gene sequences from chapter 3 as well as 2 HBV isolates, A and B (red, bold colour) rooted with one of Pacific sequences (3309). Genotype A, purple; Genotype B, blue; genotype C, green and genotype D, red colour. Indian sequences shown by bold red colour. Figure shows bootstrap value of 70% and scale donates present diversity.


0.1
U8
169
1118
144
154
7.29
1113
138
U1
159
139
0083
0081
8.22
8.32
7.13
B
A
155
8.25
0098
0062
0076
0123
0018
0094
0052
3414
5029
2007
0026
0012
A-19
8/19
A-16
U10
168
U5
U6
U4
8.31
U7
9.5
U2
U9
8.21
7.32
160
198
5072
A-50
A-23
A-9
A-5
A-15
A-45
A-11
A-12
A-14
A-42
A-38
A-43
A-48
A-34
A-20
A-10
A-1
A-44
A-55
A-13
A-36
A-35
A-3
A-4
A-8
A-2
A-24
A-54
A-21
U3
A-53
A-32
A-52
A-40
A-37
A-49
A-51
A-41
2080
5017
3415
3097
0078
3309

References:

1. **Abbas, A. K., K. M. Murphy, and A. Sher.** 1996. Functional diversity of helper T lymphocytes. Nature **383:**787-93.

2. **Abdelhamed, A. M., C. M. Kelley, T. G. Miller, P. A. Furman, and H. C. Isom.** 2002. Rebound of hepatitis B virus replication in HepG2 cells after cessation of antiviral treatment. J Virol **76:**8148-60.

3. **Ackrill, A. M., N. V. Naoumov, A. L. Eddleston, and R. Williams.** 1993. Specific deletions in the hepatitis B virus core open reading frame in patients with chronic active hepatitis B. J Med Virol **41:**165-9.

4. **Ahn, S. H., K. H. Han, J. Y. Park, C. K. Lee, S. W. Kang, C. Y. Chon, Y. S. Kim, K. Park, D. K. Kim, and Y. M. Moon.** 2000. Association between hepatitis B virus infection and HLA-DR type in Korea. Hepatology **31:**1371-3.

5. **Aiba, N., M. J. McGarvey, J. Waters, S. J. Hadziyannis, H. C. Thomas, and P. Karayiannis.** 1997. The precore sequence of hepatitis B virus is required for nuclear localization of the core protein. Hepatology **26:**1311-7.

6. **Aikawa, T., K. Kanai, M. Kako, T. Kawasaki, K. Hino, S. Iwabuchi, H. Tsubouchi, Y. Takehira, F. Tsuda, H. Okamoto, and et al.** 1995. Interferon-alpha 2a for chronic hepatitis B with e antigen or antibody:comparable antiviral effects on wild-type virus and precore mutant. J Viral Hepat **2:**243-50.

7. **Akarca, U. S., S. Greene, and A. S. Lok.** 1994. Detection of precore hepatitis B virus mutants in asymptomatic HBsAg-positive family members. Hepatology **19:**1366-70.

8. **Akarca, U. S., and A. S. Lok.** 1995b. Naturally occurring core-gene-defective hepatitis B viruses. J Gen Virol **76:**1821-6.

9. **Akarca, U. S., and A. S. Lok.** 1995a. Naturally occurring hepatitis B virus core gene mutations. Hepatology **22:**50-60.

10. **Akiba, T., H. Nakayama, Y. Miyazaki, A. Kanno, M. Ishii, and H. Ohori.** 1987. Relationship between the replication of hepatitis B virus and the localization of virus nucleocapsid antigen (HBcAg) in hepatocytes. J Gen Virol **68:**871-7.

11. **Albin, C., and W. S. Robinson.** 1980. Protein kinase activity in hepatitis B virus. J Virol **34:**297-302.

12. **Alexopoulou, A., P. Karayiannis, S. J. Hadziyannis, N. Aiba, and H. C. Thomas.** 1997. Emergence and selection of HBV variants in an anti-HBe positive patient persistently infected with quasi-species. J Hepatol **26:**748-53.

13. **Allen, M. I., M. Deslauriers, C. W. Andrews, G. A. Tipples, K. A. Walters, D. L. Tyrrell, N. Brown, and L. D. Condreay.** 1998. Identification and characterization of mutations in hepatitis B virus resistant to lamivudine. Lamivudine Clinical Investigation Group. Hepatology **27:**1670-7.

14. **Almarri, A., and J. R. Batchelor.** 1994. HLA and hepatitis B infection. Lancet **344:**1194-5.

15. **Alter, M. J.** 1996. Epidemiology and disease burden of hepatitis B and C. Antivira Thera **1 (Supp 3):**9-15.

16. **Alter, M. J.** 2003. Epidemiology of hepatitis B in Europe and worldwide. J Hepatol **39 Suppl 1:**S64-9.

17. **Alter, M. J., M. S. Favero, L. A. Moyer, J. K. Miller, and L. A. Bland.** 1990. National surveillance of dialysis-associated diseases in the United States, 1988. ASAIO Trans **36:**107-18.

18. **Andre, F. E.** 1989. Summary of safety and efficacy data on a yeast-derived hepatitis B vaccine. Am J Med **87:**14S-20S.

19. **Arauz-Ruiz, P., H. Norder, B. H. Robertson, and L. O. Magnius.** 2002. Genotype H: a new Amerindian genotype of hepatitis B virus revealed in Central America. J Gen Virol **83:**2059-73.

20. **Argos, P., and S. D. Fuller.** 1988. A model for the hepatitis B virus core protein: prediction of antigenic sites and relationship to RNA virus capsid proteins. Embo J **7:**819-24.

21. **Ballare, M., C. Lavarini, M. R. Brunetto, E. Petruzzelli, M. Dovis, G. Molino, and F. Bonino.** 1989. Relationship between the intrahepatic expression of 'e' and 'c' epitopes of the nucleocapsid protein of hepatitis B virus and viraemia. Clin Exp Immunol **75:**64-9.

22. **Baptista, M., A. Kramvis, and M. C. Kew.** 1999. High prevalence of 1762(T) 1764(A) mutations in the basic core promoter of hepatitis B virus isolated from black Africans with hepatocellular carcinoma compared with asymptomatic carriers. Hepatology **29:**946-53.

23. **Baron, J. L., L. Gardiner, S. Nishimura, K. Shinkai, R. Locksley, and D. Ganem.** 2002. Activation of a nonclassical NKT cell subset in a transgenic mouse model of hepatitis B virus infection. Immunity **16:**583-94.

24. **Bartenschlager, R., M. Junker-Niepmann, and H. Schaller.** 1990. The P gene product of hepatitis B virus is required as a structural component for genomic RNA encapsidation. J Virol **64:**5324-32.

25. **Bartholomeusz, A., and S. Locarnini.** 2001. Hepatitis B virus mutants and fulminant hepatitis B: fitness plus phenotype. Hepatology **34:**432-5.

26. **Basuni.** 2001. Aspects of HBV S gene variation. Glasgow, Glasgow.

27. **Baumert, T. F., A. Marrone, J. Vergalla, and T. J. Liang.** 1998. Naturally occurring mutations define a novel function of the hepatitis B virus core promoter in core protein expression. J Virol **72:**6785-95.

28. **Bchini, R., F. Capel, C. Dauguet, S. Dubanchet, and M. A. Petit.** 1990. In vitro infection of human hepatoma (HepG2) cells with hepatitis B virus. J Virol **64:**3025-32.

29. **Beck, J., and M. Nassal.** 1995. Efficient hammerhead ribozyme-mediated cleavage of the structured hepatitis B virus encapsidation signal in vitro and in cell extracts, but not in intact cells. Nucleic Acids Res **23:**4954-62.

30. **Beckel-Mitchener, A., and J. Summers.** 1997. A novel transcriptional element in circular DNA monomers of the duck hepatitis B virus. J Virol **71:**7917-22.

31. **Benhamou, Y., M. Bochet, V. Thibault, V. Calvez, M. H. Fievet, P. Vig, C. S. Gibbs, C. Brosgart, J. Fry, H. Namini, C. Katlama, and T. Poynard.** 2001. Safety and efficacy of adefovir dipivoxil in patients co-infected with HIV-1 and lamivudine-resistant hepatitis B virus: an open-label pilot study. Lancet **358:**718-23.

32. **Benz, J., and A. Hofmann.** 1997. Annexins: from structure to function. Biol Chem **378:**177-83.

33. **Bertoletti, A., C. Ferrari, F. Fiaccadori, A. Penna, R. Margolskee, H. J. Schlicht, P. Fowler, S. Guilhot, and F. V. Chisari.** 1991. HLA class I-restricted human cytotoxic T cells recognize endogenously synthesized hepatitis B virus nucleocapsid antigen. Proc Natl Acad Sci U S A **88:**10445-9.

34. **Blackberg, J., and K. Kidd-Ljunggren.** 2003. Mutations within the hepatitis B virus genome among chronic hepatitis B patients with hepatocellular carcinoma. J Med Virol **71:**18-23.

35. **Blackberg, J., and K. Kidd-Ljunggren.** 2000. Occult hepatitis B virus after acute self-limited infection persisting for 30 years without sequence variation. J Hepatol **33:**992-7.

36. **Bock, C. T., B. Buerke, H. L. Tillmann, F. Tacke, V. Kliem, M. P. Manns, and C. Trautwein.** 2003. Relevance of hepatitis B core gene deletions in patients after kidney transplantation. Gastroenterology **124:**1809-20.

37. **Bock, C. T., S. Schwinn, C. H. Schroder, I. Velhagen, and H. Zentgraf.** 1996. Localization of hepatitis B virus core protein and viral DNA at the nuclear membrane. Virus Genes **12:**53-63.

38. **Bock, C. T., H. L. Tillmann, M. P. Manns, and C. Trautwein.** 1999. The pre-S region determines the intracellular localization and appearance of hepatitis B virus. Hepatology **30:**517-25.

39. **Bollyky, P. L., and E. C. Holmes.** 1999. Reconstructing the complex evolutionary history of hepatitis B virus. J Mol Evol **49:**130-41.

40. **Boner, W., H. J. Schlicht, K. Hanrieder, E. C. Holmes, and W. F. Carman.** 1995. Further characterization of 2 types of precore variant hepatitis B virus isolates from Hong Kong. J Infect Dis **171:**1461-7.

41. **Bonino, F., B. Hoyer, J. Nelson, R. Engle, G. Verme, and J. Gerin.** 1981. Hepatitis B virus DNA in the sera of HBsAg carriers: a marker of active hepatitis B virus replication in the liver. Hepatology **1:**386-91.

42. **Bottcher, B., S. A. Wynne, and R. A. Crowther.** 1997. Determination of the fold of the core protein of hepatitis B virus by electron cryomicroscopy. Nature **386:**88-91.

43. **Bozkaya, H., U. S. Akarca, B. Ayola, and A. S. Lok.** 1997. High degree of conservation in the hepatitis B virus core gene during the immune tolerant phase in perinatally acquired chronic hepatitis B virus infection. J Hepatol **26:**508-16.

44. **Bozkaya, H., B. Ayola, and A. S. Lok.** 1996. High rate of mutations in the hepatitis B core gene during the immune clearance phase of chronic hepatitis B virus infection. Hepatology **24:**32-7.

45. **Brechot, C., M. Hadchouel, J. Scotto, F. Degos, P. Charnay, C. Trepo, and P. Tiollais.** 1981. Detection of hepatitis B virus DNA in liver and serum: a direct appraisal of the chronic carrier state. Lancet **2:**765-8.

46. **Brechot, C., C. Trepo, D. Ninova-Bradistilova, A. Goudeau, C. Degott, P. Maupas, F. Potet, and J. P. Benhamou.** 1980. Hepatitis B core antigen in hepatocytes of patients with chronic active hepatitis. Dig Dis Sci **25:**593-6.

47. **Bringas, R.** 1997. Folding and assembly of hepatitis B virus core protein: a new model proposal. J Struct Biol **118:**189-96.

48. **Brunetto, M. R., F. Oliveri, P. Colombatto, M. Capalbo, C. Barbera, and F. Bonino.** 1995. Treatment of chronic anti-HBe-positive hepatitis B with interferon-alpha. J Hepatol **22:**42-4.

49. **Budkowska, A., P. Dubreuil, T. Poynard, P. Marcellin, M. A. Loriot, P. Maillard, and J. Pillot.** 1992. Anti-pre-S responses and viral clearance in chronic hepatitis B virus infection. Hepatology **15:**26-31.

50. **Budkowska, A., C. Quan, F. Groh, P. Bedossa, P. Dubreuil, J. P. Bouvet, and J. Pillot.** 1993. Hepatitis B virus (HBV) binding factor in human serum: candidate for a soluble form of hepatocyte HBV receptor. J Virol **67:**4316-22.

51. **Caillat-Zucman, S., J. J. Gimenez, F. Wambergue, G. Albouze, B. Lebkiri, C. Naret, A. Moynot, P. Jungers, and J. F. Bach.** 1998. Distinct HLA class II alleles determine antibody response to vaccination with hepatitis B surface antigen. Kidney Int **53:**1626-30.

52. **Cao, T., I. Desombere, P. Vanlandschoot, M. Sallberg, and G. Leroux-Roels.** 2002. Characterization of HLA DR13-restricted CD4(+) T cell epitopes of hepatitis B core antigen associated with self-limited, acute hepatitis B. J Gen Virol **83:**3023-33.

53. **Cao, T., P. Meuleman, I. Desombere, M. Sallberg, and G. Leroux-Roels.** 2001. In vivo inhibition of anti-hepatitis B virus core antigen (HBcAg) immunoglobulin G production by HBcAg-specific CD4(+) Th1-type T-cell clones in a hu-PBL-NOD/SCID mouse model. J Virol **75:**11449-56.

54. **Carbonara, A., W. Mayr, M. Rizzetto, L. Contu, E. S. Curtoni, M. DeMarchi, P. Farci, C. Lavarini, and E. Olivetti.** 1983. Endemic HBV

infection, tissue autoantibodies and HLA. Analysis of a Sardinian population. Tissue Antigens **22:**289-94.

55. **Cariani, E., G. Fiordalisi, and D. Primi.** 1992. Hepatitis B virus (HBV) genomic variations in chronic hepatitis. Arch Virol Suppl **4:**86-9.

56. **Carlier, D., O. Jean-Jean, N. Fouillot, H. Will, and J. M. Rossignol.** 1995. Importance of the C terminus of the hepatitis B virus precore protein in secretion of HBe antigen. J Gen Virol **76:**1041-5.

57. **Carman, W., H. Thomas, and E. Domingo.** 1993. Viral genetic variation: hepatitis B virus as a clinical example. Lancet **341:**349-53.

58. **Carman, W. F.** 1996. Molecular variants of hepatitis B virus. Clin Lab Med **16:**407-28.

59. **Carman, W. F.** 1995a. Variation in the core and X genes of hepatitis B virus. Intervirology **38:**75-88.

60. **Carman, W. F., W. Boner, G. Fattovich, K. Colman, E. S. Dornan, M. Thursz, and S. Hadziyannis.** 1997b. Hepatitis B virus core protein mutations are concentrated in B cell epitopes in progressive disease and in T helper cell epitopes during clinical remission. J Infect Dis **175:**1093-100.

61. **Carman, W. F., M. R. Jacyna, S. Hadziyannis, P. Karayiannis, M. J. McGarvey, A. Makris, and H. C. Thomas.** 1989. Mutation preventing formation of hepatitis B e antigen in patients with chronic hepatitis B infection. Lancet **2:**588-91.

62. **Carman WF, M. G., Hadziyannis S, Fattovich G, Alberti A, Thomas HC.** 1994. Core protein evolution after selection of HBV pre core mutants and correlation with disease severity. Viral hepatitis and liver disease:273-276.

63. **Carman WF, M. L.** 1997a. Pre S/S gene variants of HBV. Edizioni Minerva Medica, Turin.

64. **Carman WF, T. H., Zuckerman AJ, Harrison TJ.** 1998. Molecular variants. Churchill Livingston, Hong Kong.

65. **Carman, W. F., M. Thursz, S. Hadziyannis, G. McIntyre, K. Colman, A. Gioustoz, G. Fattovich, A. Alberti, and H. C. Thomas.** 1995b. Hepatitis B e antigen negative chronic active hepatitis: hepatitis B virus core mutations occur predominantly in known antigenic determinants. J Viral Hepat **2:**77-84.

66. **Carman, W. F., A. R. Zanetti, P. Karayiannis, J. Waters, G. Manzillo, E. Tanzi, A. J. Zuckerman, and H. C. Thomas.** 1990. Vaccine-induced escape mutant of hepatitis B virus. Lancet **336:**325-9.

67. **Chan, H. L., M. Hussain, and A. S. Lok.** 1999. Different hepatitis B virus genotypes are associated with different mutations in the core promoter and precore regions during hepatitis B e antigen seroconversion. Hepatology **29:**976-84.

68. **Chan, H. L., M. L. Wong, A. Y. Hui, L. C. Hung, F. K. Chan, and J. J. Sung.** 2003. Hepatitis B virus genotype C takes a more aggressive disease course than hepatitis B virus genotype B in hepatitis B e antigen-positive patients. J Clin Microbiol **41:**1277-9.

69. **Chayama, K., Y. Suzuki, M. Kobayashi, A. Tsubota, M. Hashimoto, Y. Miyano, H. Koike, I. Koida, Y. Arase, S. Saitoh, N. Murashima, K. Ikeda, and H. Kumada.** 1998. Emergence and takeover of YMDD motif mutant hepatitis B virus during long-term lamivudine therapy and re-takeover by wild type after cessation of therapy. Hepatology **27:**1711-6.

70. **Chen, M., M. Sallberg, S. N. Thung, J. Hughes, J. Jones, and D. R. Milich.** 2001. Modeling the T-helper cell response in acute and chronic hepatitis B virus infection using T-cell receptor transgenic mice. Antiviral Res **52:**99-111.

71. **Chen, Y., W. S. Robinson, and P. L. Marion.** 1992. Naturally occurring point mutation in the C terminus of the polymerase gene prevents duck hepatitis B virus RNA packaging. J Virol **66:**1282-7.

72. **Chen, Y., W. S. Robinson, and P. L. Marion.** 1994. Selected mutations of the duck hepatitis B virus P gene RNase H domain affect both RNA packaging and priming of minus-strand DNA synthesis. J Virol **68:**5232-8.

73. **Chisari, F. V.** 1997. Immunobiology and pathogenesis of viral hepatitis. Edizioni Minerva Medica, Turin.

74. **Chisari, F. V., and C. Ferrari.** 1995. Hepatitis B virus immunopathology. Springer Semin Immunopathol **17:**261-81.

75. **Chisari, F. V., P. Filippi, J. Buras, A. McLachlan, H. Popper, C. A. Pinkert, R. D. Palmiter, and R. L. Brinster.** 1987. Structural and pathological

effects of synthesis of hepatitis B virus large envelope polypeptide in transgenic mice. Proc Natl Acad Sci U S A **84:**6909-13.

76. **Cho SW, H. K., Kim JH.** 2000. Reversion from pre core/core promoter mutants to wild type HBV during the course of lamivudine therapy. Hepatology **32:**1163-9.

77. **Chu, C. J., E. B. Keeffe, S. H. Han, R. P. Perrillo, A. D. Min, C. Soldevila-Pico, W. Carey, R. S. Brown, Jr., V. A. Luketic, N. Terrault, and A. S. Lok.** 2003b. Hepatitis B virus genotypes in the United States: results of a nationwide study. Gastroenterology **125:**444-51.

78. **Chu, C. J., E. B. Keeffe, S. H. Han, R. P. Perrillo, A. D. Min, C. Soldevila-Pico, W. Carey, R. S. Brown, Jr., V. A. Luketic, N. Terrault, and A. S. Lok.** 2003a. Prevalence of HBV precore/core promoter variants in the United States. Hepatology **38:**619-28.

79. **Chu, C. M., and Y. F. Liaw.** 1987. Intrahepatic distribution of hepatitis B surface and core antigens in chronic hepatitis B virus infection. Hepatocyte with cytoplasmic/membranous hepatitis B core antigen as a possible target for immune hepatocytolysis. Gastroenterology **92:**220-5.

80. **Chu, C. M., and Y. F. Liaw.** 1995b. Membrane staining for hepatitis B surface antigen on hepatocytes: a sensitive and specific marker of active viral replication in hepatitis B. J Clin Pathol **48:**470-3.

81. **Chu, C. M., W. C. Shyu, R. W. Kuo, and Y. F. Liaw.** 1988. HLA class I antigen display on hepatocyte membrane in chronic hepatitis B virus infection: its role in the pathogenesis of chronic type B hepatitis. Hepatology **8:**712-7.

82. **Chu, C. M., C. T. Yeh, R. N. Chien, I. S. Sheen, and Y. F. Liaw.** 1997. The degrees of hepatocyte nuclear but not cytoplasmic expression of hepatitis B core antigen reflect the level of viral replication in chronic hepatitis B virus infection. J Clin Microbiol **35:**102-5.

83. **Chu, C. M., C. T. Yeh, I. S. Sheen, and Y. F. Liaw.** 1995a. Subcellular localization of hepatitis B core antigen in relation to hepatocyte regeneration in chronic hepatitis B. Gastroenterology **109:**1926-32.

84. **Chuang, W. L., M. Omata, T. Ehata, O. Yokosuka, and M. Ohto.** 1993. Concentrating missense mutations in core gene of hepatitis B virus. Evidence for adaptive mutation in chronic hepatitis B virus infection. Dig Dis Sci **38:**594-600.

85. **CIA** 1999, posting date. .HTTP://WWW.odci.gov/cia/publications/factbook. [Online.]

86. **Clements, M. L., E. Miskovsky, M. Davidson, T. Cupps, N. Kumwenda, L. A. Sandman, D. West, T. Hesley, V. Ioli, W. Miller, and et al.** 1994. Effect of age on the immunogenicity of yeast recombinant hepatitis B vaccines containing surface antigen (S) or PreS2 + S antigens. J Infect Dis **170:**510-6.

87. **Colucci, G., Y. Beazer, C. Cantaluppi, and C. Tackney.** 1988. Identification of a major hepatitis B core antigen (HBcAg) determinant by using synthetic peptides and monoclonal antibodies. J Immunol **141:**4376-80.

88. **Conway, J. F., N. Cheng, A. Zlotnick, S. J. Stahl, P. T. Wingfield, D. M. Belnap, U. Kanngiesser, M. Noah, and A. C. Steven.** 1998b. Hepatitis B virus capsid: localization of the putative immunodominant loop (residues 78 to 83) on the capsid surface, and implications for the distinction between c and e-antigens. J Mol Biol **279:**1111-21.

89. **Conway, J. F., N. Cheng, A. Zlotnick, S. J. Stahl, P. T. Wingfield, and A. C. Steven.** 1998a. Localization of the N terminus of hepatitis B virus capsid protein by peptide-based difference mapping from cryoelectron microscopy. Proc Natl Acad Sci U S A **95:**14622-7.

90. **Cooper, A., N. Paran, and Y. Shaul.** 2003. The earliest steps in hepatitis B virus infection. Biochim Biophys Acta **1614:**89-96.

91. **Courouce-Pauty, A. M., J. M. Lemaire, and J. F. Roux.** 1978. New hepatitis B surface antigen subtypes inside the ad category. Vox Sang **35:**304-8.

92. **Courouce-Pauty, A. M., A. Plancon, and J. P. Soulier.** 1983. Distribution of HBsAg subtypes in the world. Vox Sang **44:**197-211.

93. **Crowther, R. A., N. A. Kiselev, B. Bottcher, J. A. Berriman, G. P. Borisova, V. Ose, and P. Pumpens.** 1994. Three-dimensional structure of hepatitis B virus core particles determined by electron cryomicroscopy. Cell **77:**943-50.

94. **Cui, C., J. Shi, L. Hui, H. Xi, Zhuoma, Quni, Tsedan, and G. Hu.** 2002. The dominant hepatitis B virus genotype identified in Tibet is a C/D hybrid. J Gen Virol **83:**2773-7.

95. **Cupps, T. R., J. Tibbles, W. M. Hurni, W. J. Miller, R. W. Ellis, D. Milich, and N. Wetter.** 1993. In vitro T cell immune responses to the preS2 antigen of the hepatitis B virus envelope protein in preS2 + S vaccine recipients. Absence of cross-reactivity of subtypes at a major T cell recognition site. J Immunol **151:**3353-60.

96. **Dane, D. S., C. H. Cameron, and M. Briggs.** 1970. Virus-like particles in serum of patients with Australia-antigen-associated hepatitis. Lancet **1:**695-8.

97. **Dash, S., K. V. Rao, and S. K. Panda.** 1992. Receptor for pre-S1(21-47) component of hepatitis B virus on the liver cell: role in virus cell interaction. J Med Virol **37:**116-21.

98. **Daub, H., S. Blencke, P. Habenberger, A. Kurtenbach, J. Dennenmoser, J. Wissing, A. Ullrich, and M. Cotten.** 2002. Identification of SRPK1 and SRPK2 as the major cellular protein kinases phosphorylating hepatitis B virus core protein. J Virol **76:**8124-37.

99. **Davis, H. L., Suparto, II, R. R. Weeratna, Jumintarto, D. D. Iskandriati, S. S. Chamzah, A. A. Ma'ruf, C. C. Nente, D. D. Pawitri, A. M. Krieg, Heriyanto, W. Smits, and D. D. Sajuthi.** 2000. CpG DNA overcomes hyporesponsiveness to hepatitis B vaccine in orangutans. Vaccine **18:**1920-4.

100. **de Bruin, W., W. Leenders, T. Kos, K. Hertogs, E. Depla, and S. H. Yap.** 1994. Hepatitis delta virus attaches to human hepatocytes via human liver endonexin II, a specific HBsAg binding protein. J Viral Hepat **1:**33-8.

101. **de Bruin, W. C., K. Hertogs, W. P. Leenders, E. Depla, and S. H. Yap.** 1995. Hepatitis B virus: specific binding and internalization of small HBsAg by human hepatocytes. J Gen Virol **76 (Pt 4):**1047-50.

102. **de Bruin, W. C., W. P. Leenders, H. Moshage, and U. J. van Haelst.** 1996. Species specificity for HBsAg binding protein endonexin II. J Hepatol **24:**265-70.

103. **De Falco, S., M. G. Ruvoletto, A. Verdoliva, M. Ruvo, A. Raucci, M. Marino, S. Senatore, G. Cassani, A. Alberti, P. Pontisso, and G. Fassina.** 2001. Cloning and expression of a novel hepatitis B virus-binding protein from HepG2 cells. J Biol Chem **276:**36613-23.

104. **De Meyer, S., Z. J. Gong, W. Suwandhi, J. van Pelt, A. Soumillion, and S. H. Yap.** 1997. Organ and species specificity of hepatitis B virus (HBV)

infection: a review of literature with a special reference to preferential attachment of HBV to human hepatocytes. J Viral Hepat **4:**145-53.

105. **Diamond.** 1988. Expression train to Polynesia. Nature **336:**307-308.

106. **Dienstag, J. L., R. P. Perrillo, E. R. Schiff, M. Bartholomew, C. Vicary, and M. Rubin.** 1995. A preliminary trial of lamivudine for chronic hepatitis B infection. N Engl J Med **333:**1657-61.

107. **Dienstag, J. L., E. R. Schiff, T. L. Wright, R. P. Perrillo, H. W. Hann, Z. Goodman, L. Crowther, L. D. Condreay, M. Woessner, M. Rubin, and N. A. Brown.** 1999. Lamivudine as initial treatment for chronic hepatitis B in the United States. N Engl J Med **341:**1256-63.

108. **Diepolder, H. M., M. C. Jung, E. Keller, W. Schraut, J. T. Gerlach, N. Gruner, R. Zachoval, R. M. Hoffmann, C. A. Schirren, S. Scholz, and G. R. Pape.** 1998. A vigorous virus-specific CD4+ T cell response may contribute to the association of HLA-DR13 with viral clearance in hepatitis B. Clin Exp Immunol **113:**244-51.

109. **Ding, X., H. Gu, Z. H. Zhong, X. Zilong, H. T. Tran, Y. Iwaki, T. C. Li, T. Sata, and K. Abe.** 2003. Molecular epidemiology of hepatitis viruses and genotypic distribution of hepatitis B and C viruses in Harbin, China. Jpn J Infect Dis **56:**19-22.

110. **Domingo, E., A. Mas, E. Yuste, N. Pariente, S. Sierra, M. Gutierrez-Riva, and L. Menendez-Arias.** 2001. Virus population dynamics, fitness variations and the control of viral disease: an update. Prog Drug Res **57:**77-115.

111. **Dyson, M. R., and K. Murray.** 1995. Selection of peptide inhibitors of interactions involved in complex protein assemblies: association of the core and surface antigens of hepatitis B virus. Proc Natl Acad Sci U S A **92:**2194-8.

112. **Eckhardt, S. G., D. R. Milich, and A. McLachlan.** 1991. Hepatitis B virus core antigen has two nuclear localization sequences in the arginine-rich carboxyl terminus. J Virol **65:**575-82.

113. **Ehata, T., M. Omata, W. L. Chuang, O. Yokosuka, Y. Ito, K. Hosoda, and M. Ohto.** 1993. Mutations in core nucleotide sequence of hepatitis B virus correlate with fulminant and severe hepatitis. J Clin Invest **91:**1206-13.

114. **Ehata, T., M. Omata, O. Yokosuka, K. Hosoda, and M. Ohto.** 1991. Amino acid residues of core region of hepatitis B virus. Asymptomatic carriers versus patients with liver disease. J Gastroenterol Hepatol **6:**292-6.

115. **Ehata, T., M. Omata, O. Yokosuka, K. Hosoda, and M. Ohto.** 1992. Variations in codons 84-101 in the core nucleotide sequence correlate with hepatocellular injury in chronic hepatitis B virus infection. J Clin Invest **89:**332-8.

116. **Ehata, T., Omata M, Chuang WL, Yokosuka O, Ito Y, Hosoda K, Ohto M.** 1994. Mutations in the core nucleotide sequence of HBV correlate with severe liver disease. Viral hepatitis and liver disease:277-282.

117. **Erhardt, A., U. Reineke, D. Blondin, W. H. Gerlich, O. Adams, T. Heintges, C. Niederau, and D. Haussinger.** 2000. Mutations of the core promoter and response to interferon treatment in chronic replicative hepatitis B. Hepatology **31:**716-25.

118. **Esteban, J., Martell M, Carman WF, Gomez J.** 1999. The impact of rapid evolution of the hepatitis viruses. Academic press, London.

119. **Fang, J. W., P. C. Wu, C. L. Lai, C. K. Lo, A. Meager, and J. Y. Lau.** 1994. Hepatic expression of interferon-alpha in chronic hepatitis B virus infection. Dig Dis Sci **39:**2014-21.

120. **Farza, H., M. Hadchouel, J. Scotto, P. Tiollais, C. Babinet, and C. Pourcel.** 1988. Replication and gene expression of hepatitis B virus in a transgenic mouse that contains the complete viral genome. J Virol **62:**4144-52.

121. **Fattovich, G., L. Brollo, S. Boscaro, P. Pontisso, G. Giustina, D. Criscuolo, D. Maladorno, A. Alberti, G. Realdi, and A. Ruol.** 1989. Long-term effect of low dose recombinant interferon therapy in patients with chronic hepatitis B. J Hepatol **9:**331-7.

122. **Feitelson, M. A., M. Zhu, L. X. Duan, and W. T. London.** 1993. Hepatitis B x antigen and p53 are associated in vitro and in liver tissues from patients with primary hepatocellular carcinoma. Oncogene **8:**1109-17.

123. **Feray, C., M. Gigou, D. Samuel, J. Bernuau, H. Bismuth, and C. Brechot.** 1993. Low prevalence of precore mutations in hepatitis B virus DNA in fulminant hepatitis type B in France. J Hepatol **18:**119-22.

124. **Ferns, R. B., and R. S. Tedder.** 1986. Human and monoclonal antibodies to hepatitis B core antigen recognise a single immunodominant epitope. J Med Virol **19:**193-203.

125. **Ferns, R. B., and R. S. Tedder.** 1984. Monoclonal antibodies to hepatitis Be antigen (HBeAg) derived from hepatitis B core antigen (HBcAg): their use in characterization and detection of HBeAg. J Gen Virol **65:**899-908.

126. **Ferrari, C., A. Bertoletti, A. Penna, A. Cavalli, A. Valli, G. Missale, M. Pilli, P. Fowler, T. Giuberti, F. V. Chisari, and et al.** 1991. Identification of immunodominant T cell epitopes of the hepatitis B virus nucleocapsid antigen. J Clin Invest **88:**214-22.

127. **Ferrari, C., A. Penna, A. Bertoletti, A. Cavalli, A. Valli, G. Missale, M. Pilli, S. Marchelli, T. Giuberti, and F. Fiaccadori.** 1992. Immune pathogenesis of hepatitis B. Arch Virol Suppl **4:**11-8.

128. **Ferrari, C., A. Penna, A. Bertoletti, A. Valli, A. D. Antoni, T. Giuberti, A. Cavalli, M. A. Petit, and F. Fiaccadori.** 1990. Cellular immune response to hepatitis B virus-encoded antigens in acute and chronic hepatitis B virus infection. J Immunol **145:**3442-9.

129. **Ferrari, C., A. Penna, T. Giuberti, M. J. Tong, E. Ribera, F. Fiaccadori, and F. V. Chisari.** 1987. Intrahepatic, nucleocapsid antigen-specific T cells in chronic active hepatitis B. J Immunol **139:**2050-8.

130. **Franco, A., M. Paroli, U. Testa, R. Benvenuto, C. Peschle, F. Balsano, and V. Barnaba.** 1992. Transferrin receptor mediates uptake and presentation of hepatitis B envelope antigen by T lymphocytes. J Exp Med **175:**1195-205.

131. **Friedt, M., P. Gerner, E. Lausch, H. Trubel, B. Zabel, and S. Wirth.** 1999. Mutations in the basic core promotor and the precore region of hepatitis B virus and their selection in children with fulminant and chronic hepatitis B. Hepatology **29:**1252-8.

132. **Fujiwara, K., O. Yokosuka, T. Ehata, W. L. Chuang, F. Imazeki, H. Saisho, and M. Omata.** 1998. The two different states of hepatitis B virus DNA in asymptomatic carriers: HBe-antigen-positive versus anti-HBe-positive asymptomatic carriers. Dig Dis Sci **43:**368-76.

133. **Galle, P. R., J. Hagelstein, B. Kommerell, M. Volkmann, P. Schranz, and H. Zentgraf.** 1994. In vitro experimental infection of primary human hepatocytes with hepatitis B virus. Gastroenterology **106:**664-73.

134. **Gallimore, A., A. Glithero, A. Godkin, A. C. Tissot, A. Pluckthun, T. Elliott, H. Hengartner, and R. Zinkernagel.** 1998. Induction and exhaustion of lymphocytic choriomeningitis virus-specific cytotoxic T lymphocytes visualized using soluble tetrameric major histocompatibility complex class I-peptide complexes. J Exp Med **187:**1383-93.

135. **Gandhe, S. S., M. S. Chadha, A. M. Walimbe, and V. A. Arankalle.** 2003. Hepatitis B virus: prevalence of precore/core promoter mutants in different clinical categories of Indian patients. J Viral Hepat **10:**367-82.

136. **Ganem, D.** 1996. Hepadnaviridae and their replication, vol. 2. Lippincott-Raven, philadelphia.

137. **Ganem, D., J. R. Pollack, and J. Tavis.** 1994. Hepatitis B virus reverse transcriptase and its many roles in hepadnaviral genomic replication. Infect Agents Dis **3:**85-93.

138. **Ganem, D., and H. E. Varmus.** 1987. The molecular biology of the hepatitis B viruses. Annu Rev Biochem **56:**651-93.

139. **Garcia, P. D., J. H. Ou, W. J. Rutter, and P. Walter.** 1988. Targeting of the hepatitis B virus precore protein to the endoplasmic reticulum membrane: after signal peptide cleavage translocation can be aborted and the product released into the cytoplasm. J Cell Biol **106:**1093-104.

140. **Gazina, E. V., J. E. Fielding, B. Lin, and D. A. Anderson.** 2000. Core protein phosphorylation modulates pregenomic RNA encapsidation to different extents in human and duck hepatitis B viruses. J Virol **74:**4721-8.

141. **Geissler, M., K. Tokushige, C. C. Chante, V. R. Zurawski, Jr., and J. R. Wands.** 1997. Cellular and humoral immune response to hepatitis B virus structural proteins in mice after DNA-based immunization. Gastroenterology **112:**1307-20.

142. **Gerlich, W. H., U. Goldmann, R. Muller, W. Stibbe, and W. Wolff.** 1982. Specificity and localization of the hepatitis B virus-associated protein kinase. J Virol **42:**761-6.

143. **Gerlich, W. H., and W. S. Robinson.** 1980. Hepatitis B virus contains protein attached to the 5' terminus of its complete DNA strand. Cell **21:**801-9.

144. **Gerner, P. R., M. Friedt, R. Oettinger, E. Lausch, and S. Wirth.** 1998. The hepatitis B virus seroconversion to anti-HBe is frequently associated with HBV genotype changes and selection of preS2-defective particles in chronically infected children. Virology **245:**163-72.

145. **Gitlin, N.** 1997. Hepatitis B: diagnosis, prevention, and treatment. Clin Chem **43:**1500-6.

146. **Gong, Y., E. Yao, M. Stevens, and J. E. Tavis.** 2000. Evidence that the first strand-transfer reaction of duck hepatitis B virus reverse transcription requires the polymerase and that strand transfer is not needed for the switch of the polymerase to the elongation mode of DNA synthesis. J Gen Virol **81:**2059-65.

147. **Gong, Z. J., S. De Meyer, J. van Pelt, K. Hertogs, E. Depla, A. Soumillion, J. Fevery, and S. H. Yap.** 1999. Transfection of a rat hepatoma cell line with a construct expressing human liver annexin V confers susceptibility to hepatitis B virus infection. Hepatology **29:**576-84.

148. **Goodbourn, S., L. Didcock, and R. E. Randall.** 2000. Interferons: cell signalling, immune modulation, antiviral response and virus countermeasures. J Gen Virol **81:**2341-64.

149. **Gowans, E. J., C. J. Burrell, A. R. Jilbert, and B. P. Marmion.** 1983. Patterns of single- and double-stranded hepatitis B virus DNA and viral antigen accumulation in infected liver cells. J Gen Virol **64:**1229-39.

150. **Grady, G. F., V. A. Lee, A. M. Prince, G. L. Gitnick, K. A. Fawaz, G. N. Vyas, M. D. Levitt, J. R. Senior, J. T. Galambos, T. E. Bynum, J. W. Singleton, B. F. Clowdus, K. Akdamar, R. D. Aach, E. I. Winkelman, G. M. Schiff, and T. Hersh.** 1978. Hepatitis B immune globulin for accidental exposures among medical personnel: final report of a multicenter controlled trial. J Infect Dis **138:**625-38.

151. **Greenspan, D., P. Palese, and M. Krystal.** 1988. Two nuclear location signals in the influenza virus NS1 nonstructural protein. J Virol **62:**3020-6.

152. **Gripon, P., S. Rumin, S. Urban, J. Le Seyec, D. Glaise, I. Cannie, C. Guyomard, J. Lucas, C. Trepo, and C. Guguen-Guillouzo.** 2002. Infection of a human hepatoma cell line by hepatitis B virus. Proc Natl Acad Sci U S A **99:**15655-60.

153. **Guidotti, L. G.** 2002. The role of cytotoxic T cells and cytokines in the control of hepatitis B virus infection. Vaccine **20 Suppl 4:**A80-2.

154. **Guidotti, L. G., and F. V. Chisari.** 2001. Noncytolytic control of viral infections by the innate and adaptive immune response. Annu Rev Immunol **19:**65-91.

155. **Guidotti, L. G., S. Guilhot, and F. V. Chisari.** 1994. Interleukin-2 and alpha/beta interferon down-regulate hepatitis B virus gene expression in vivo by tumor necrosis factor-dependent and - independent pathways. J Virol **68:**1265-70.

156. **Guidotti, L. G., T. Ishikawa, M. V. Hobbs, B. Matzke, R. Schreiber, and F. V. Chisari.** 1996. Intracellular inactivation of the hepatitis B virus by cytotoxic T lymphocytes. Immunity **4:**25-36.

157. **Guidotti, L. G., V. Martinez, Y. T. Loh, C. E. Rogler, and F. V. Chisari.** 1994. Hepatitis B virus nucleocapsid particles do not cross the hepatocyte nuclear membrane in transgenic mice. J Virol **68:**5469-75.

158. **Guidotti, L. G., B. Matzke, and F. V. Chisari.** 1997. Hepatitis B virus replication is cell cycle independent during liver regeneration in transgenic mice. J Virol **71:**4804-8.

159. **Guidotti, L. G., R. Rochford, J. Chung, M. Shapiro, R. Purcell, and F. V. Chisari.** 1999. Viral clearance without destruction of infected cells during acute HBV infection. Science **284:**825-9.

160. **Gunther, S., L. Fischer, I. Pult, M. Sterneck, and H. Will.** 1999. Naturally occurring variants of hepatitis B virus. Adv Virus Res **52:**25-137.

161. **Gunther, S., H. Meisel, A. Reip, S. Miska, D. H. Kruger, and H. Will.** 1992. Frequent and rapid emergence of mutated pre-C sequences in HBV from e-antigen positive carriers who seroconvert to anti-HBe during interferon treatment. Virology **187:**271-9.

162. **Gunther, S., N. Piwon, A. Iwanska, R. Schilling, H. Meisel, and H. Will.** 1996. Type, prevalence, and significance of core promoter/enhancer II mutations in hepatitis B viruses from immunosuppressed patients with severe liver disease. J Virol **70:**8318-31.

163. **Gunther, S., G. Sommer, F. Von Breunig, A. Iwanska, T. Kalinina, M. Sterneck, and H. Will.** 1998. Amplification of full-length hepatitis B virus genomes from samples from patients with low levels of viremia: frequency

and functional consequences of PCR-introduced mutations. J Clin Microbiol **36:**531-8.

164. **Hadziyannis, S., G. Raimondo, C. Papaioannou, C. Anastassakos, D. Wong, J. Sninsky, and D. Shafritz.** 1987. Expression of pre-S gene-encoded proteins in liver and serum during chronic hepatitis B virus infection in comparison to other markers of active virus replication. J Hepatol **5:**253-9.

165. **Hadziyannis, S. J., H. M. Lieberman, G. G. Karvountzis, and D. A. Shafritz.** 1983. Analysis of liver disease, nuclear HBcAg, viral replication, and hepatitis B virus DNA in liver and serum of HBeAg Vs. anti-HBe positive carriers of hepatitis B virus. Hepatology **3:**656-62.

166. **Hagelberg, E., and J. B. Clegg.** 1993. Genetic polymorphisms in prehistoric Pacific islanders determined by analysis of ancient bone DNA. Proc R Soc Lond B Biol Sci **252:**163-70.

167. **Hanks, S. K., A. M. Quinn, and T. Hunter.** 1988. The protein kinase family: conserved features and deduced phylogeny of the catalytic domains. Science **241:**42-52.

168. **Hannoun, C., P. Horal, and M. Lindh.** 2000a. Long-term mutation rates in the hepatitis B virus genome. J Gen Virol **81:**75-83.

169. **Hannoun, C., H. Norder, and M. Lindh.** 2000b. An aberrant genotype revealed in recombinant hepatitis B virus strains from Vietnam. J Gen Virol **81:**2267-72.

170. **Hasegawa, K., J. Huang, S. A. Rogers, H. E. Blum, and T. J. Liang.** 1994. Enhanced replication of a hepatitis B virus mutant associated with an epidemic of fulminant hepatitis. J Virol **68:**1651-9.

171. **Hawkins, A. E., M. A. Zuckerman, M. Briggs, R. J. Gilson, A. H. Goldstone, N. S. Brink, and R. S. Tedder.** 1996. Hepatitis B nucleotide sequence analysis: linking an outbreak of acute hepatitis B to contamination of a cryopreservation tank. J Virol Methods **60:**81-8.

172. **Heijtink, R. A., G. A. De Wilde, J. Kruining, L. Berk, J. Balzarini, E. De Clercq, A. Holy, and S. W. Schalm.** 1993. Inhibitory effect of 9-(2-phosphonylmethoxyethyl)-adenine (PMEA) on human and duck hepatitis B virus infection. Antiviral Res **21:**141-53.

173. **Heineman T. C, C.-M., M L., Poland G.A., Jacobson, RM., Izu A. E., Sakamoto D.,.** 1999. A randomized, controlled study in adults of the

immunogenicity of a novel hepatitis B vaccine containing MF59 adjuvant. Vaccine **17:**2769-78.

174. **Hillis, W. D., A. Hillis, W. B. Bias, and W. G. Walker.** 1977. Associations of hepatitis B surface antigenemia with HLA locus B specificities. N Engl J Med **296:**1310-4.

175. **Hohler, T., G. Gerken, A. Notghi, R. Lubjuhn, H. Taheri, U. Protzer, H. F. Lohr, P. M. Schneider, K. H. Meyer zum Buschenfelde, and C. Rittner.** 1997. HLA-DRB1*1301 and *1302 protect against chronic hepatitis B. J Hepatol **26:**503-7.

176. **Hohler, T., B. Stradmann-Bellinghausen, R. Starke, R. Sanger, A. Victor, C. Rittner, and P. M. Schneider.** 2002. C4A deficiency and nonresponse to hepatitis B vaccination. J Hepatol **37:**387-92.

177. **Holland, J., and E. Domingo.** 1998. Origin and evolution of viruses. Virus Genes **16:**13-21.

178. **Honigwachs, J., O. Faktor, R. Dikstein, Y. Shaul, and O. Laub.** 1989. Liver-specific expression of hepatitis B virus is determined by the combined action of the core gene promoter and the enhancer. J Virol **63:**919-24.

179. **Hoofnagle, J. H.** 1990. Alpha-interferon therapy of chronic hepatitis B. Current status and recommendations. J Hepatol **11:**S100-7.

180. **Hoofnagle, J. H., and A. M. di Bisceglie.** 1997. The treatment of chronic viral hepatitis. N Engl J Med **336:**347-56.

181. **Horwitz, M. S., T. Krahl, C. Fine, J. Lee, and N. Sarvetnick.** 1999. Protection from lethal coxsackievirus-induced pancreatitis by expression of gamma interferon. J Virol **73:**1756-66.

182. **Hosono, S., P. C. Tai, W. Wang, M. Ambrose, D. G. Hwang, T. T. Yuan, B. H. Peng, C. S. Yang, C. S. Lee, and C. Shih.** 1995. Core antigen mutations of human hepatitis B virus in hepatomas accumulate in MHC class II-restricted T cell epitopes. Virology **212:**151-62.

183. **Hou, J., G. K. Lau, J. Cheng, C. C. Cheng, K. Luo, and W. F. Carman.** 1999. T1762/A1764 variants of the basal core promoter of hepatitis B virus; serological and clinical correlations in Chinese patients. Liver **19:**411-7.

184. **Hsu, H. C., I. J. Su, M. Y. Lai, D. S. Chen, M. H. Chang, S. M. Chuang, and J. L. Sung.** 1987. Biologic and prognostic significance of hepatocyte hepatitis B core antigen expressions in the natural course of chronic hepatitis B virus infection. J Hepatol **5:**45-50.

185. **Huang, M., and J. Summers.** 1994. pet, a small sequence distal to the pregenome cap site, is required for expression of the duck hepatitis B virus pregenome. J Virol **68:**1564-72.

186. **Huang, S. N., and A. R. Neurath.** 1979. Immunohistologic demonstration of hepatitis B viral antigens in liver with reference to its significance in liver injury. Lab Invest **40:**1-17.

187. **Hui, E. K., K. L. Chen, and S. J. Lo.** 1999. Hepatitis B virus maturation is affected by the incorporation of core proteins having a C-terminal substitution of arginine or lysine stretches. J Gen Virol **80:**2661-71.

188. **Hultgren, C., D. R. Milich, O. Weiland, and M. Sallberg.** 1998. The antiviral compound ribavirin modulates the T helper (Th) 1/Th2 subset balance in hepatitis B and C virus-specific immune responses. J Gen Virol **79:**2381-91.

189. **Huovila, A. P., A. M. Eder, and S. D. Fuller.** 1992. Hepatitis B surface antigen assembles in a post-ER, pre-Golgi compartment. J Cell Biol **118:**1305-20.

190. **Hur, G. M., Y. I. Lee, D. J. Suh, and J. H. Lee.** 1996. Gradual accumulation of mutations in precore core region of HBV in patients with chronic active hepatitis: implications of clustering changes in a small region of the HBV core region. J Med Virol **48:**38-46.

191. **Hwang, G. Y., C. J. Huang, C. Y. Lin, and C. C. Wu.** 2003. Dominant mutations of hepatitis B virus variants in hepatoma accumulate in B-cell and T-cell epitopes of the HBx antigen. Virus Res **92:**157-64.

192. **Hwang, Y. K., N. K. Kim, J. M. Park, K. Lee, W. K. Han, H. I. Kim, and H. S. Cheong.** 2002. HLA-A2 1 restricted peptides from the HBx antigen induce specific CTL responses in vitro and in vivo. Vaccine **20:**3770-7.

193. **Hyodo, N., M. Tajimi, T. Ugajin, I. Nakamura, and M. Imawari.** 2003. Frequencies of interferon-gamma and interleukin-10 secreting cells in peripheral blood mononuclear cells and liver infiltrating lymphocytes in chronic hepatitis B virus infection. Hepatol Res **27:**109-116.

194. **Inoue, K., M. Yoshiba, K. Sekiyama, H. Okamoto, and M. Mayumi.** 1998. Clinical and molecular virological differences between fulminant hepatic failures following acute and chronic infection with hepatitis B virus. J Med Virol **55:**35-41.

195. **Jiang, Y. G., Y. M. Wang, T. H. Liu, and J. Liu.** 2003. Association between HLA class II gene and susceptibility or resistance to chronic hepatitis B. World J Gastroenterol **9:**2221-5.

196. **Jones, C. D., M. Page, A. Bacon, E. Cahill, M. Bentley, and S. N. Chatfield.** 1998. Characterization of the T- and B-cell immune response to a new recombinant pre-S1, pre-S2 and SHBs antigen containing hepatitis B vaccine (Hepagene); evidence for superior anti-SHBs antibody induction in responder mice. J Viral Hepat **5 Suppl 2:**5-8.

197. **Jung, M. C., H. M. Diepolder, U. Spengler, E. A. Wierenga, R. Zachoval, R. M. Hoffmann, D. Eichenlaub, G. Frosner, H. Will, and G. R. Pape.** 1995. Activation of a heterogeneous hepatitis B (HB) core and e antigen-specific CD4+ T-cell population during seroconversion to anti-HBe and anti-HBs in hepatitis B virus infection. J Virol **69:**3358-68.

198. **Kakimi, K., L. G. Guidotti, Y. Koezuka, and F. V. Chisari.** 2000. Natural killer T cell activation inhibits hepatitis B virus replication in vivo. J Exp Med **192:**921-30.

199. **Kakimi, K., T. E. Lane, F. V. Chisari, and L. G. Guidotti.** 2001. Cutting edge: Inhibition of hepatitis B virus replication by activated NK T cells does not require inflammatory cell recruitment to the liver. J Immunol **167:**6701-5.

200. **Kakumu, S., M. Arao, K. Yoshioka, Y. Tsutsumi, and M. Inoue.** 1989. Distribution of HBcAg in hepatitis B detected by immunoperoxidase staining with three different preparations of anti-HBc antibodies. J Clin Pathol **42:**284-8.

201. **Kakumu, S., T. Ishikawa, M. Mizokami, E. Orido, K. Yoshioka, T. Wakita, and M. Yamamoto.** 1991. Treatment with human gamma interferon of chronic hepatitis B: comparative study with alpha interferon. J Med Virol **35:**32-7.

202. **Kalderon, D., B. L. Roberts, W. D. Richardson, and A. E. Smith.** 1984. A short amino acid sequence able to specify nuclear location. Cell **39:**499-509.

203. **Kamimura, T., A. Yoshikawa, F. Ichida, and H. Sasaki.** 1981. Electron microscopic studies of Dane particles in hepatocytes with special reference to intracellular development of Dane particles and their relation with HBeAg in serum. Hepatology **1:**392-7.

204. **Kane, M. A.** 1998. Global status of HB immunization, 1998. Acta Gastroenterol Belg **61:**237.

205. **Kann, M., and W. H. Gerlich.** 1994. Effect of core protein phosphorylation by protein kinase C on encapsidation of RNA within core particles of hepatitis B virus. J Virol **68:**7993-8000.

206. **Kann M, G. W.** 1998. Structure and molecular biology. Churchill Livingston, Hong Kong.

207. **Kann, M., X. Lu, and W. H. Gerlich.** 1995. Recent studies on replication of hepatitis B virus. J Hepatol **22(1:**9-13.

208. **Kann, M., B. Sodeik, A. Vlachou, W. H. Gerlich, and A. Helenius.** 1999. Phosphorylation-dependent binding of hepatitis B virus core particles to the nuclear pore complex. J Cell Biol **145:**45-55.

209. **Kao, J. H., P. J. Chen, M. Y. Lai, and D. S. Chen.** 2001. Acute exacerbations of chronic hepatitis B are rarely associated with superinfection of hepatitis B virus. Hepatology **34:**817-23.

210. **Kao, J. H., P. J. Chen, M. Y. Lai, and D. S. Chen.** 2003. Basal core promoter mutations of hepatitis B virus increase the risk of hepatocellular carcinoma in hepatitis B carriers. Gastroenterology **124:**327-34.

211. **Kao, J. H., P. J. Chen, M. Y. Lai, and D. S. Chen.** 2000a. Hepatitis B genotypes correlate with clinical outcomes in patients with chronic hepatitis B. Gastroenterology **118:**554-9.

212. **Kao, J. H., C. J. Liu, and D. S. Chen.** 2002. Hepatitis B viral genotypes and lamivudine resistance. J Hepatol **36:**303-4.

213. **Kao, J. H., N. H. Wu, P. J. Chen, M. Y. Lai, and D. S. Chen.** 2000b. Hepatitis B genotypes and the response to interferon therapy. J Hepatol **33:**998-1002.

214. **Karayiannis, P.** 2003. Hepatitis B virus: old, new and future approaches to antiviral treatment. J Antimicrob Chemother **51:**761-85.

215. **Kawai, K., N. Horiike, K. Michitaka, and M. Onji.** 2003. The effects of hepatitis B virus core promoter mutations on hepatitis B core antigen distribution in hepatocytes as detected by laser-assisted microdissection. J Hepatol **38:**635-41.

216. **Kayser, M., S. Brauer, G. Weiss, P. A. Underhill, L. Roewer, W. Schiefenhovel, and M. Stoneking.** 2000. Melanesian origin of Polynesian Y chromosomes. Curr Biol **10:**1237-46.

217. **Khakoo, S. I., R. Ling, I. Scott, A. I. Dodi, T. J. Harrison, G. M. Dusheiko, and J. A. Madrigal.** 2000. Cytotoxic T lymphocyte responses and CTL epitope escape mutation in HBsAg, anti-HBe positive individuals. Gut **47:**137-43.

218. **Kidd-Ljunggren, K., K. Ekdahl, M. Oberg, S. Kurathong, and S. Lolekha.** 1995. Hepatitis B virus strains in Thailand: genomic variants in chronic carriers. J Med Virol **47:**454-61.

219. **Kidd-Ljunggren, K., Y. Miyakawa, and A. H. Kidd.** 2002. Genetic variability in hepatitis B viruses. J Gen Virol **83:**1267-80.

220. **Kidd-Ljunggren, K., and O. Simonsen.** 1999. Reappearance of hepatitis B 10 years after kidney transplantation. N Engl J Med **341:**127-8.

221. **Kim, Y. K., E. Junn, I. Park, Y. Lee, C. Kang, and J. K. Ahn.** 1999. Repression of hepatitis B virus X gene expression by hammerhead ribozymes. Biochem Biophys Res Commun **257:**759-65.

222. **Kimura, K., K. Kakimi, S. Wieland, L. G. Guidotti, and F. V. Chisari.** 2002. Activated intrahepatic antigen-presenting cells inhibit hepatitis B virus replication in the liver of transgenic mice. J Immunol **169:**5188-95.

223. **Kimura, K., K. Kakimi, S. Wieland, L. G. Guidotti, and F. V. Chisari.** 2002. Interleukin-18 inhibits hepatitis B virus replication in the livers of transgenic mice. J Virol **76:**10702-7.

224. **Klenerman, P., Y. Wu, and R. Phillips.** 2002. HIV: current opinion in escapology. Curr Opin Microbiol **5:**408-13.

225. **Kock, J., and H. J. Schlicht.** 1993. Analysis of the earliest steps of hepadnavirus replication: genome repair after infectious entry into hepatocytes does not depend on viral polymerase activity. J Virol **67:**4867-74.

226. **Kodama, K., N. Ogasawara, H. Yoshikawa, and S. Murakami.** 1985. Nucleotide sequence of a cloned woodchuck hepatitis virus genome: evolutional relationship between hepadnaviruses. J Virol **56:**978-86.

227. **Koibuchi, T., A. Hitani, T. Nakamura, N. Nojiri, K. Nakajima, T. Jyuji, and A. Iwamoto.** 2001. Predominance of genotype A HBV in an HBV-HIV-1 dually positive population compared with an HIV-1-negative counterpart in Japan. J Med Virol **64:**435-40.

228. **Kratz, P. A., B. Bottcher, and M. Nassal.** 1999. Native display of complete foreign protein domains on the surface of hepatitis B virus capsids. Proc Natl Acad Sci U S A **96:**1915-20.

229. **Kuroki, K., R. Cheung, P. L. Marion, and D. Ganem.** 1994. A cell surface protein that binds avian hepatitis B virus particles. J Virol **68:**2091-6.

230. **Lai, C. L., R. N. Chien, N. W. Leung, T. T. Chang, R. Guan, D. I. Tai, K. Y. Ng, P. C. Wu, J. C. Dent, J. Barber, S. L. Stephenson, and D. F. Gray.** 1998. A one-year trial of lamivudine for chronic hepatitis B. Asia Hepatitis Lamivudine Study Group. N Engl J Med **339:**61-8.

231. **Lai, C. L., C. K. Ching, A. K. Tung, E. Li, J. Young, A. Hill, B. C. Wong, J. Dent, and P. C. Wu.** 1997. Lamivudine is effective in suppressing hepatitis B virus DNA in Chinese hepatitis B surface antigen carriers: a placebo-controlled trial. Hepatology **25:**241-4.

232. **Lai, C. L., A. S. Lok, H. J. Lin, P. C. Wu, E. K. Yeoh, and C. Y. Yeung.** 1987. Placebo-controlled trial of recombinant alpha 2-interferon in Chinese HBsAg-carrier children. Lancet **2:**877-80.

233. **Lakhdar, M., A. Senik, and W. H. Fridman.** 1984. Human cytotoxic T lymphocytes (CTL) against Epstein-Barr virus (EBV) infected cells: EBV specificity and involvement of major histocompatibility complex determinants in the lysis exerted by anti- EBV CTL toward HLA-compatible and allogeneic target cells. Cell Immunol **83:**414-21.

234. **Lan, Y. T., J. Li, W. Liao, and J. Ou.** 1999. Roles of the three major phosphorylation sites of hepatitis B virus core protein in viral replication. Virology **259:**342-8.

235. **Lanford, R. E., D. Chavez, K. M. Brasky, R. B. Burns, 3rd, and R. Rico-Hesse.** 1998. Isolation of a hepadnavirus from the woolly monkey, a New World primate. Proc Natl Acad Sci U S A **95:**5757-61.

236. **Laras, A., J. Koskinas, K. Avgidis, and S. J. Hadziyannis.** 1998. Incidence and clinical significance of hepatitis B virus precore gene translation initiation mutations in e antigen-negative patients. J Viral Hepat **5:**241-8.

237. **Laras, A., J. Koskinas, and S. J. Hadziyannis.** 2002. In vivo suppression of precore mRNA synthesis is associated with mutations in the hepatitis B virus core promoter. Virology **295:**86-96.

238. **Laskus, T., D. H. Persing, M. J. Nowicki, J. W. Mosley, and J. Rakela.** 1993. Nucleotide sequence analysis of the precore region in patients with fulminant hepatitis B in the United States. Gastroenterology **105:**1173-8.

239. **Lau, j. Y. N., Alexander GJM., Alberti A.** 1991. Viral hepatitis. Gut **5:**S47-S62.

240. **Lautenschlager, I., E. Taskinen, K. Inkinen, V. P. Lehto, I. Virtanen, and P. Hayry.** 1984. Distribution of the major histocompatibility complex antigens on different cellular components of human liver. Cell Immunol **85:**191-200.

241. **Lazdina, U., M. Alheim, J. Nystrom, C. Hultgren, G. Borisova, I. Sominskaya, P. Pumpens, D. L. Peterson, D. R. Milich, and M. Sallberg.** 2003. Priming of cytotoxic T cell responses to exogenous hepatitis B virus core antigen is B cell dependent. J Gen Virol **84:**139-46.

242. **Lazdina, U., T. Cao, J. Steinbergs, M. Alheim, P. Pumpens, D. L. Peterson, D. R. Milich, G. Leroux-Roels, and M. Sallberg.** 2001. Molecular basis for the interaction of the hepatitis B virus core antigen with the surface immunoglobulin receptor on naive B cells. J Virol **75:**6367-74.

243. **Lechner, F., D. K. Wong, P. R. Dunbar, R. Chapman, R. T. Chung, P. Dohrenwend, G. Robbins, R. Phillips, P. Klenerman, and B. D. Walker.** 2000. Analysis of successful immune responses in persons infected with hepatitis C virus. J Exp Med **191:**1499-512.

244. **Lee, P. I., L. Y. Chang, C. Y. Lee, L. M. Huang, and M. H. Chang.** 1997. Detection of hepatitis B surface gene mutation in carrier children with or without immunoprophylaxis at birth. J Infect Dis **176:**427-30.

245. **Lee, W.** 1997. Hepatitis B virus infection. N Engl J Med **337:**1733-1745.

246. **Lee, Y. I., G. M. Hur, D. J. Suh, and S. H. Kim.** 1996. Novel pre-C/C gene mutants of hepatitis B virus in chronic active hepatitis: naturally occurring escape mutants. J Gen Virol **77:**1129-38.

247. **Lenhoff, R. J., C. A. Luscombe, and J. Summers.** 1998. Competition in vivo between a cytopathic variant and a wild-type duck hepatitis B virus. Virology **251:**85-95.

248. **Lenhoff, R. J., and J. Summers.** 1994. Coordinate regulation of replication and virus assembly by the large envelope protein of an avian hepadnavirus. J Virol **68:**4565-71.

249. **Li, J. S., S. P. Tong, and J. R. Wands.** 1996. Characterization of a 120-Kilodalton pre-S-binding protein as a candidate duck hepatitis B virus receptor. J Virol **70:**6029-35.

250. **Li, J. S., S. P. Tong, Y. M. Wen, L. Vitvitski, Q. Zhang, and C. Trepo.** 1993. Hepatitis B virus genotype A rarely circulates as an HBe-minus mutant: possible contribution of a single nucleotide in the precore region. J Virol **67:**5402-10.

251. **Liao, W., and J. H. Ou.** 1995. Phosphorylation and nuclear localization of the hepatitis B virus core protein: significance of serine in the three repeated SPRRR motifs. J Virol **69:**1025-9.

252. **Liaw, Y. F., R. N. Chien, C. T. Yeh, S. L. Tsai, and C. M. Chu.** 1999. Acute exacerbation and hepatitis B virus clearance after emergence of YMDD motif mutation during lamivudine therapy. Hepatology **30:**567-72.

253. **Liaw, Y. F., N. W. Leung, T. T. Chang, R. Guan, D. I. Tai, K. Y. Ng, R. N. Chien, J. Dent, L. Roman, S. Edmundson, and C. L. Lai.** 2000. Effects of extended lamivudine therapy in Asian patients with chronic hepatitis B. Asia Hepatitis Lamivudine Study Group. Gastroenterology **119:**172-80.

254. **Lien, J. M., C. E. Aldrich, and W. S. Mason.** 1986. Evidence that a capped oligoribonucleotide is the primer for duck hepatitis B virus plus-strand DNA synthesis. J Virol **57:**229-36.

255. **Lindh, M., C. Hannoun, A. P. Dhillon, G. Norkrans, and P. Horal.** 1999. Core promoter mutations and genotypes in relation to viral replication and liver damage in East Asian hepatitis B virus carriers. J Infect Dis **179:**775-82.

256. **Liu, C. J., J. H. Kao, W. Y. Shau, P. J. Chen, M. Y. Lai, and D. S. Chen.** 2002. Naturally occurring hepatitis B surface gene variants in chronic hepatitis B virus infection: correlation with viral serotypes and clinical stages of liver disease. J Med Virol **68:**50-9.

257. **Loeb, D. D., K. J. Gulya, and R. Tian.** 1997. Sequence identity of the terminal redundancies on the minus-strand DNA template is necessary but not sufficient for the template switch during hepadnavirus plus-strand DNA synthesis. J Virol **71:**152-60.

258. **Loeb, D. D., R. C. Hirsch, and D. Ganem.** 1991. Sequence-independent RNA cleavages generate the primers for plus strand DNA synthesis in hepatitis B viruses: implications for other reverse transcribing elements. Embo J **10:**3533-40.

259. **Loffler-Mary, H., J. Dumortier, C. Klentsch-Zimmer, and R. Prange.** 2000. Hepatitis B virus assembly is sensitive to changes in the cytosolic S loop of the envelope proteins. Virology **270:**358-67.

260. **Lok A S, B. H., Ayola B.** 1997. HBV pre core/core gene mutants: epidemiology. Edizioni Minerva Medica, Turin.

261. **Lok, A. S., C. L. Lai, P. C. Wu, and E. K. Leung.** 1988. Long-term follow-up in a randomised controlled trial of recombinant alpha 2-interferon in Chinese patients with chronic hepatitis B infection. Lancet **2:**298-302.

262. **Lu, C. C., and T. S. Yen.** 1996. Activation of the hepatitis B virus S promoter by transcription factor NF-Y via a CCAAT element. Virology **225:**387-94.

263. **Lu, M., G. Hilken, J. Kruppenbacher, T. Kemper, R. Schirmbeck, J. Reimann, and M. Roggendorf.** 1999. Immunization of woodchucks with plasmids expressing woodchuck hepatitis virus (WHV) core antigen and surface antigen suppresses WHV infection. J Virol **73:**281-9.

264. **Lucas, M., U. Karrer, A. Lucas, and P. Klenerman.** 2001. Viral escape mechanisms--escapology taught by viruses. Int J Exp Pathol **82:**269-86.

265. **Lyons, R. H., B. Q. Ferguson, and M. Rosenberg.** 1987. Pentapeptide nuclear localization signal in adenovirus E1a. Mol Cell Biol **7:**2451-6.

266. **MacDonald, D. M., E. C. Holmes, J. C. Lewis, and P. Simmonds.** 2000. Detection of hepatitis B virus infection in wild-born chimpanzees (Pan troglodytes verus): phylogenetic relationships with human and other primate genotypes. J Virol **74:**4253-7.

267. **Machida, A., H. Ohnuma, E. Takai, F. Tsuda, T. Tanaka, M. Naito, E. Munekata, Y. Miyakawa, and M. Mayumi.** 1989. Antigenic sites on the arginine-rich carboxyl-terminal domain of the capsid protein of hepatitis B virus distinct from hepatitis B core or e antigen. Mol Immunol **26:**413-21.

268. **Machida, A., H. Ohnuma, F. Tsuda, A. Yoshikawa, Y. Hoshi, T. Tanaka, S. Kishimoto, Y. Akahane, Y. Miyakawa, and M. Mayumi.** 1991. Phosphorylation in the carboxyl-terminal domain of the capsid protein of hepatitis B virus: evaluation with a monoclonal antibody. J Virol **65:**6024-30.

269. **Maddrey.** 2000. Hepatitis B: an important public health issue. J Med Virol **61:**362-6.

270. **Magnius, L. O., and H. Norder.** 1995. Subtypes, genotypes and molecular epidemiology of the hepatitis B virus as reflected by sequence variability of the S-gene. Intervirology **38:**24-34.

271. **Malanchere-Bres, E., P. J. Payette, M. Mancini, P. Tiollais, H. L. Davis, and M. L. Michel.** 2001. CpG oligodeoxynucleotides with hepatitis B surface antigen (HBsAg) for vaccination in HBsAg-transgenic mice. J Virol **75:**6482-91.

272. **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:**782-92.

273. **Marinos, G., H. M. Smith, N. V. Naoumov, and R. Williams.** 1994. Quantitative assessment of serum IgM anti-HBc in the natural course and during interferon treatment of chronic hepatitis B virus infection. Hepatology **19:**303-11.

274. **Marinos, G., F. Torre, S. Gunther, M. G. Thomas, H. Will, R. Williams, and N. V. Naoumov.** 1996. Hepatitis B virus variants with core gene deletions in the evolution of chronic hepatitis B infection. Gastroenterology **111:**183-92.

275. **Maruyama, T., S. Kuwata, K. Koike, S. Iino, K. Yasuda, H. Yotsuyanagi, K. Moriya, H. Maekawa, H. Yamada, Y. Shibata, and D. R.**

Milich. 1998. Precore wild-type DNA and immune complexes persist in chronic hepatitis B after seroconversion: no association between genome conversion and seroconversion. Hepatology **27:**245-53.

276. **Mayerat, C., A. Mantegani, and P. C. Frei.** 1999. Does hepatitis B virus (HBV) genotype influence the clinical outcome of HBV infection? J Viral Hepat **6:**299-304.

277. **Maynard, J. E., M. A. Kane, and S. C. Hadler.** 1989. Global control of hepatitis B through vaccination: role of hepatitis B vaccine in the Expanded Programme on Immunization. Rev Infect Dis **11 Suppl 3:**S574-8.

278. **McAleer, W. J., E. B. Buynak, R. Z. Maigetter, D. E. Wampler, W. J. Miller, and M. R. Hilleman.** 1984. Human hepatitis B vaccine from recombinant yeast. Nature **307:**178-80.

279. **McMichael, A. J., A. Ting, H. J. Zweerink, and B. A. Askonas.** 1977. HLA restriction of cell-mediated lysis of influenza virus-infected human cells. Nature **270:**524-6.

280. **McMillan, J. S., D. S. Bowden, P. W. Angus, G. W. McCaughan, and S. A. Locarnini.** 1996. Mutations in the hepatitis B virus precore/core gene and core promoter in patients with severe recurrent disease following liver transplantation. Hepatology **24:**1371-8.

281. **Mehdi, H., M. J. Kaplan, F. Y. Anlar, X. Yang, R. Bayer, K. Sutherland, and M. E. Peeples.** 1994. Hepatitis B virus surface antigen binds to apolipoprotein H. J Virol **68:**2415-24.

282. **Mehdi, H., X. Yang, and M. E. Peeples.** 1996. An altered form of apolipoprotein H binds hepatitis B virus surface antigen most efficiently. Virology **217:**58-66.

283. **Milich, D. R.** 1997c. Influence of T-helper cell subsets and crossregulation in hepatitis B virus infection. J Viral Hepat **4:**48-59.

284. **Milich, D. R., M. Chen, F. Schodel, D. L. Peterson, J. E. Jones, and J. L. Hughes.** 1997a. Role of B cells in antigen presentation of the hepatitis B core. Proc Natl Acad Sci U S A **94:**14648-53.

285. **Milich, D. R., M. K. Chen, J. L. Hughes, and J. E. Jones.** 1998. The secreted hepatitis B precore antigen can modulate the immune response to the nucleocapsid: a mechanism for persistence. J Immunol **160:**2013-21.

286. **Milich, D. R., J. L. Hughes, R. Houghten, A. McLachlan, and J. E. Jones.** 1989. Functional identification of agretopic and epitopic residues within an HBcAg T cell determinant. J Immunol **143:**3141-7.

287. **Milich, D. R., J. L. Hughes, A. McLachlan, K. E. Langley, G. B. Thornton, and J. E. Jones.** 1990. Importance of subtype in the immune response to the pre-S(2) region of the hepatitis B surface antigen. I. T cell fine specificity. J Immunol **144:**3535-43.

288. **Milich, D. R., and G. G. Leroux-Roels.** 2003. Immunogenetics of the response to HBsAg vaccination. Autoimmun Rev **2:**248-57.

289. **Milich, D. R., and A. McLachlan.** 1986. The nucleocapsid of hepatitis B virus is both a T-cell-independent and a T-cell-dependent antigen. Science **234:**1398-401.

290. **Milich, D. R., A. McLachlan, S. Stahl, P. Wingfield, G. B. Thornton, J. L. Hughes, and J. E. Jones.** 1988. Comparative immunogenicity of hepatitis B virus core and E antigens. J Immunol **141:**3617-24.

291. **Milich, D. R., A. McLachlan, G. B. Thornton, and J. L. Hughes.** 1987. Antibody production to the nucleocapsid and envelope of the hepatitis B virus primed by a single synthetic T cell site. Nature **329:**547-9.

292. **Milich, D. R., D. L. Peterson, F. Schodel, J. E. Jones, and J. L. Hughes.** 1995b. Preferential recognition of hepatitis B nucleocapsid antigens by Th1 or Th2 cells is epitope and major histocompatibility complex dependent. J Virol **69:**2776-85.

293. **Milich, D. R., M. Sallberg, and T. Maruyama.** 1995a. The humoral immune response in acute and chronic hepatitis B virus infection. Springer Semin Immunopathol **17:**149-66.

294. **Milich, D. R., F. Schodel, J. L. Hughes, J. E. Jones, and D. L. Peterson.** 1997b. The hepatitis B virus core and e antigens elicit different Th cell subsets: antigen structure can affect Th cell phenotype. J Virol **71:**2192-201.

295. **Miller, R. H., and W. S. Robinson.** 1986. Common evolutionary origin of hepatitis B virus and retroviruses. Proc Natl Acad Sci U S A **83:**2531-5.

296. **Missale, G., A. Redeker, J. Person, P. Fowler, S. Guilhot, H. J. Schlicht, C. Ferrari, and F. V. Chisari.** 1993. HLA-A31- and HLA-Aw68-

restricted cytotoxic T cell responses to a single hepatitis B virus nucleocapsid epitope during acute viral hepatitis. J Exp Med **177:**751-62.

297. **Mizokami, M., and E. Orito.** 1999. [Molecular evolution of hepatitis viruses]. Nippon Shokakibyo Gakkai Zasshi **96:**1033-43.

298. **Mizokami, M., E. Orito, K. Ohba, K. Ikeo, J. Y. Lau, and T. Gojobori.** 1997. Constrained evolution with respect to gene overlap of hepatitis B virus. J Mol Evol **44:**S83-90.

299. **Mohiti, J., A. M. Caswell, and J. H. Walker.** 1997. The nuclear location of annexin V in the human osteosarcoma cell line MG-63 depends on serum factors and tyrosine kinase signaling pathways. Exp Cell Res **234:**98-104.

300. **Mondelli, M., G. M. Vergani, A. Alberti, D. Vergani, B. Portmann, A. L. Eddleston, and R. Williams.** 1982. Specificity of T lymphocyte cytotoxicity to autologous hepatocytes in chronic hepatitis B virus infection: evidence that T cells are directed against HBV core antigen expressed on hepatocytes. J Immunol **129:**2773-8.

301. **Montano, L., G. C. Miescher, A. H. Goodall, K. H. Wiedmann, G. Janossy, and H. C. Thomas.** 1982. Hepatitis B virus and HLA antigen display in the liver during chronic hepatitis B virus infection. Hepatology **2:**557-61.

302. **Moreland, R. B., G. L. Langevin, R. H. Singer, R. L. Garcea, and L. M. Hereford.** 1987. Amino acid sequences that determine the nuclear localization of yeast histone 2B. Mol Cell Biol **7:**4048-57.

303. **Moriya, K., M. Matsukura, K. Kurokawa, and K. Koike.** 1996. In vivo inhibition of hepatitis B virus gene expression by antisense phosphorothioate oligonucleotides. Biochem Biophys Res Commun **218:**217-23.

304. **Morozov, V., M. Pisareva, and M. Groudinin.** 2000. Homologous recombination between different genotypes of hepatitis B virus. Gene **260:**55-65.

305. **Muller, R., G. Gubernatis, M. Farle, G. Niehoff, H. Klein, C. Wittekind, G. Tusch, H. U. Lautz, K. Boker, W. Stangel, and et al.** 1991. Liver transplantation in HBs antigen (HBsAg) carriers. Prevention of hepatitis B virus (HBV) recurrence by passive immunization. J Hepatol **13:**90-6.

306. **Mutchnick, M. G., K. L. Lindsay, E. R. Schiff, G. D. Cummings, H. D. Appelman, R. R. Peleman, M. Silva, K. C. Roach, F. Simmons, S. Milstein, S. C. Gordon, and M. N. Ehrinpreis.** 1999. Thymosin alpha1 treatment of chronic hepatitis B: results of a phase III multicentre, randomized, double-blind and placebo-controlled study. J Viral Hepat **6:**397-403.

307. **Mutimer, D., B. H. Feraz-Neto, R. Harrison, K. O'Donnell, J. Shaw, P. Cane, and D. Pillay.** 2001. Acute liver graft failure due to emergence of lamivudine resistant hepatitis B virus: rapid resolution during treatment with adefovir. Gut **49:**860-3.

308. **Nakayoshi, T., T. Maeshiro, H. Nakasone, H. Sakugawa, F. Kinjo, E. Orito, and M. Mizokami.** 2003. Difference in prognosis between patients infected with hepatitis B virus with genotype B and those with genotype C in the Okinawa Islands: a prospective study. J Med Virol **70:**350-4.

309. **Nakazono, K., Y. Ito, C. H. Wu, and G. Y. Wu.** 1996. Inhibition of hepatitis B virus replication by targeted pretreatment of complexed antisense DNA in vitro. Hepatology **23:**1297-303.

310. **Nakopoulou, L., N. Adraskelas, K. Stefanaki, D. Zacharoulis, and S. Hadziyannis.** 1992. Expression of HBsAg and HBcAg in liver tissue: correlation with disease activity. Histol Histopathol **7:**493-9.

311. **Naoumov, N. V., K. A. Antonov, S. Miska, V. Bichko, R. Williams, and H. Will.** 1997. Differentiation of core gene products of the hepatitis B virus in infected liver tissue using monoclonal antibodies. J Med Virol **53:**127-38.

312. **Naoumov, N. V., H. M. Daniels, F. Davison, A. L. Eddleston, G. J. Alexander, and R. Williams.** 1993. Identification of hepatitis B virus-DNA in the liver by in situ hybridization using a biotinylated probe. Relation to HBcAg expression and histology. J Hepatol **19:**204-10.

313. **Naoumov, N. V., R. Schneider, T. Grotzinger, M. C. Jung, S. Miska, G. R. Pape, and H. Will.** 1992. Precore mutant hepatitis B virus infection and liver disease. Gastroenterology **102:**538-43.

314. **Naoumov, N. V., M. G. Thomas, A. L. Mason, S. Chokshi, C. J. Bodicky, F. Farzaneh, R. Williams, and R. P. Perrillo.** 1995. Genomic variations in the hepatitis B core gene: a possible factor influencing response to interferon alfa treatment. Gastroenterology **108:**505-14.

315. **Nassal, M.** 1992b. The arginine-rich domain of the hepatitis B virus core protein is required for pregenome encapsidation and productive viral positive- strand DNA synthesis but not for virus assembly. J Virol **66:**4107-16.

316. **Nassal, M., and A. Rieger.** 1993. An intramolecular disulfide bridge between Cys-7 and Cys61 determines the structure of the secretory core gene product (e antigen) of hepatitis B virus. J Virol **67:**4307-15.

317. **Nassal, M., A. Rieger, and O. Steinau.** 1992a. Topological analysis of the hepatitis B virus core particle by cysteine-cysteine cross-linking. J Mol Biol **225:**1013-25.

318. **Nayersina, R., P. Fowler, S. Guilhot, G. Missale, A. Cerny, H. J. Schlicht, A. Vitiello, R. Chesnut, J. L. Person, A. G. Redeker, and et al.** 1993. HLA A2 restricted cytotoxic T lymphocyte responses to multiple hepatitis B surface antigen epitopes during hepatitis B virus infection. J Immunol **150:**4659-71.

319. **Neurath, A. R., and N. Strick.** 1994. The putative cell receptors for hepatitis B virus (HBV), annexin V, and apolipoprotein H, bind to lipid components of HBV. Virology **204:**475-7.

320. **Neurath, A. R., N. Strick, and P. Sproul.** 1992. Search for hepatitis B virus cell receptors reveals binding sites for interleukin 6 on the virus envelope protein. J Exp Med **175:**461-9.

321. **Ni, Y. H., M. H. Chang, H. Y. Hsu, and H. L. Chen.** 2000. Long-term follow-up study of core gene deletion mutants in children with chronic hepatitis B virus infection. Hepatology **32:**124-8.

322. **Nicoll, A. J., D. L. Colledge, J. J. Toole, P. W. Angus, R. A. Smallwood, and S. A. Locarnini.** 1998. Inhibition of duck hepatitis B virus replication by 9-(2-phosphonylmethoxyethyl)adenine, an acyclic phosphonate nucleoside analogue. Antimicrob Agents Chemother **42:**3130-5.

323. **Norder, H., P. Arauz-Ruiz, L. Blitz, F. H. Pujol, J. M. Echevarria, and L. O. Magnius.** 2003. The T(1858) variant predisposing to the precore stop mutation correlates with one of two major genotype F hepatitis B virus clades. J Gen Virol **84:**2083-7.

324. **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.

325. **Norder, H., A. M. Courouce, and L. O. Magnius.** 1993b. Complete nucleotide sequences of six hepatitis B viral genomes encoding the surface antigen subtypes ayw4, adw4q-, and adrq- and their phylogenetic classification. Arch Virol Suppl **8:**189-99.

326. **Norder, H., A. M. Courouce, and L. O. Magnius.** 1992b. Molecular basis of hepatitis B virus serotype variations within the four major subtypes. J Gen Virol **73:**3141-5.

327. **Norder, H., J. W. Ebert, H. A. Fields, I. K. Mushahwar, and L. O. Magnius.** 1996. Complete sequencing of a gibbon hepatitis B virus genome reveals a unique genotype distantly related to the chimpanzee hepatitis B virus. Virology **218:**214-23.

328. **Norder, H., B. Hammas, S. D. Lee, K. Bile, A. M. Courouce, I. K. Mushahwar, and L. O. Magnius.** 1993a. Genetic relatedness of hepatitis B viral strains of diverse geographical origin and natural variations in the primary structure of the surface antigen. J Gen Virol **74 (Pt 7):**1341-8.

329. **Norder, H., B. Hammas, S. Lofdahl, A. M. Courouce, and L. O. Magnius.** 1992a. Comparison of the amino acid sequences of nine different serotypes of hepatitis B surface antigen and genomic classification of the corresponding hepatitis B virus strains. J Gen Virol **73:**1201-8.

330. **Offensperger, W. B., S. Offensperger, E. Walter, H. E. Blum, and W. Gerok.** 1991. Inhibition of duck hepatitis B virus infection by lysosomotropic agents. Virology **183:**415-8.

331. **Ogata, N., R. H. Miller, K. G. Ishak, and R. H. Purcell.** 1993. The complete nucleotide sequence of a pre-core mutant of hepatitis B virus implicated in fulminant hepatitis and its biological characterization in chimpanzees. Virology **194:**263-76.

332. **Ogawa, M., K. Hasegawa, T. Naritomi, N. Torii, and N. Hayashi.** 2002. Clinical features and viral sequences of various genotypes of hepatitis B virus compared among patients with acute hepatitis B. Hepatol Res **23:**167-177.

333. **O'Grady, J. G., H. M. Smith, S. E. Davies, H. M. Daniels, P. T. Donaldson, K. C. Tan, B. Portmann, G. J. Alexander, and R. Williams.**

1992. Hepatitis B virus reinfection after orthotopic liver transplantation. Serological and clinical implications. J Hepatol **14:**104-11.

334. **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:**231-6.

335. **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:**2305-14.

336. **Okamoto H, M. M., Mayum M.** 1997. Mutations in the HBV core promoter for a decreased expression of HBeAg. Edizioni Minerva Medica, Turin.

337. **Okamoto, H., F. Tsuda, Y. Akahane, Y. Sugai, M. Yoshiba, K. Moriyama, T. Tanaka, Y. Miyakawa, and M. Mayumi.** 1994. Hepatitis B virus with mutations in the core promoter for an e antigen-negative phenotype in carriers with antibody to e antigen. J Virol **68:**8102-10.

338. **Okamoto, H., F. Tsuda, H. Sakugawa, R. I. Sastrosoewignjo, M. Imai, Y. Miyakawa, and M. Mayumi.** 1988. Typing hepatitis B virus by homology in nucleotide sequence: comparison of surface antigen subtypes. J Gen Virol **69:**2575-83.

339. **Okumura, A., M. Takayanagi, T. Aiyama, K. Iwata, T. Wakita, T. Ishikawa, K. Yoshioka, and S. Kakumu.** 1996. Serial analysis of hepatitis B virus core nucleotide sequence of patients with acute exacerbation during chronic infection. J Med Virol **49:**103-109.

340. **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:**1747-57.

341. **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 U S A **86:**7059-62.

342. **Orito, E., M. Mizokami, H. Sakugawa, K. Michitaka, K. Ishikawa, T. Ichida, T. Okanoue, H. Yotsuyanagi, and S. Iino.** 2001. A case-control

study for clinical and molecular biological differences between hepatitis B viruses of genotypes B and C. Japan HBV Genotype Research Group. Hepatology **33:**218-23.

343. **Ou, J. H., and K. D. Bell.** 1990. Comparative studies of hepatitis B virus precore and core particles. Virology **174:**185-91.

344. **Ou, J. H., O. Laub, and W. J. Rutter.** 1986. Hepatitis B virus gene function: the precore region targets the core antigen to cellular membranes and causes the secretion of the e antigen. Proc Natl Acad Sci U S A **83:**1578-82.

345. **Ou, J. H., C. T. Yeh, and T. S. Yen.** 1989. Transport of hepatitis B virus precore protein into the nucleus after cleavage of its signal peptide. J Virol **63:**5238-43.

346. **Pante, N., and U. Aebi.** 1993. The nuclear pore complex. J Cell Biol **122:**977-84.

347. **Papatheodoridis, G. V., and S. J. Hadziyannis.** 2001. Diagnosis and management of pre-core mutant chronic hepatitis B. J Viral Hepat **8:**311-21.

348. **Paran, N., B. Geiger, and Y. Shaul.** 2001. HBV infection of cell culture: evidence for multivalent and cooperative attachment. Embo J **20:**4443-53.

349. **Park, Y. N., K. H. Han, K. S. Kim, J. P. Chung, S. Kim, and C. Park.** 1999. Cytoplasmic expression of hepatitis B core antigen in chronic hepatitis B virus infection: role of precore stop mutants. Liver **19:**199-205.

350. **Parra, B., D. R. Hinton, N. W. Marten, C. C. Bergmann, M. T. Lin, C. S. Yang, and S. A. Stohlman.** 1999. IFN-gamma is required for viral clearance from central nervous system oligodendroglia. J Immunol **162:**1641-7.

351. **Patzer, E. J., G. R. Nakamura, C. C. Simonsen, A. D. Levinson, and R. Brands.** 1986. Intracellular assembly and packaging of hepatitis B surface antigen particles occur in the endoplasmic reticulum. J Virol **58:**884-92.

352. **Penna, A., F. V. Chisari, A. Bertoletti, G. Missale, P. Fowler, T. Giuberti, F. Fiaccadori, and C. Ferrari.** 1991. Cytotoxic T lymphocytes recognize an HLA-A2-restricted epitope within the hepatitis B virus nucleocapsid antigen. J Exp Med **174:**1565-70.

353. **Penna, A., G. Del Prete, A. Cavalli, A. Bertoletti, M. M. D'Elios, R. Sorrentino, M. D'Amato, C. Boni, M. Pilli, F. Fiaccadori, and C. Ferrari.** 1997. Predominant T-helper 1 cytokine profile of hepatitis B virus nucleocapsid-specific T cells in acute self-limited hepatitis B. Hepatology **25:**1022-7.

354. **Penna, A., P. Fowler, A. Bertoletti, S. Guilhot, B. Moss, R. F. Margolskee, A. Cavalli, A. Valli, F. Fiaccadori, F. V. Chisari, and et al.** 1992. Hepatitis B virus (HBV)-specific cytotoxic T-cell (CTL) response in humans: characterization of HLA class II-restricted CTLs that recognize endogenously synthesized HBV envelope antigens. J Virol **66:**1193-8.

355. **Perrillo, R. P., C. R. Campbell, S. Strang, C. J. Bodicky, and D. J. Costigan.** 1984. Immune globulin and hepatitis B immune globulin. Prophylactic measures for intimate contacts exposed to acute type B hepatitis. Arch Intern Med **144:**81-5.

356. **Perrillo, R. P., and A. L. Mason.** 1993. Hepatitis B and liver transplantation. Problems and promises. N Engl J Med **329:**1885-7.

357. **Petit, M. A., and J. Pillot.** 1985. HBc and HBe antigenicity and DNA-binding activity of major core protein P22 in hepatitis B virus core particles isolated from the cytoplasm of human liver cells. J Virol **53:**543-51.

358. **Pircher, H., D. Moskophidis, U. Rohrer, K. Burki, H. Hengartner, and R. M. Zinkernagel.** 1990. Viral escape by selection of cytotoxic T cell-resistant virus variants in vivo. Nature **346:**629-33.

359. **Pollack, J. R., and D. Ganem.** 1993. An RNA stem-loop structure directs hepatitis B virus genomic RNA encapsidation. J Virol **67:**3254-63.

360. **Pontisso, P., M. G. Ruvoletto, C. Tiribelli, W. H. Gerlich, A. Ruol, and A. Alberti.** 1992. The preS1 domain of hepatitis B virus and IgA cross-react in their binding to the hepatocyte surface. J Gen Virol **73 (Pt 8):**2041-5.

361. **Porres, J. C., V. Carreno, I. Mora, J. Gutiez, A. Moreno, S. Ramon y Cajal, J. A. Marron, and J. Bartolome.** 1988. Different doses of recombinant alpha interferon in the treatment of chronic hepatitis B patients without antibodies against the human immunodeficiency virus. Hepatogastroenterology **35:**300-3.

362. **Prange, R., and R. E. Streeck.** 1995. Novel transmembrane topology of the hepatitis B virus envelope proteins. Embo J **14:**247-56.

363. **Preikschat, P., S. Gunther, S. Reinhold, H. Will, K. Budde, H. H. Neumayer, D. H. Kruger, and H. Meisel.** 2002. Complex HBV populations with mutations in core promoter, C gene, and pre-S region are associated with development of cirrhosis in long-term renal transplant recipients. Hepatology **35:**466-77.

364. **Protzer-Knolle, U., U. Naumann, R. Bartenschlager, T. Berg, U. Hopf, K. H. Meyer zum Buschenfelde, P. Neuhaus, and G. Gerken.** 1998. Hepatitis B virus with antigenically altered hepatitis B surface antigen is selected by high-dose hepatitis B immune globulin after liver transplantation. Hepatology **27:**254-63.

365. **Pugh, J., A. Zweidler, and J. Summers.** 1989. Characterization of the major duck hepatitis B virus core particle protein. J Virol **63:**1371-6.

366. **Pushko, P., M. Sallberg, G. Borisova, U. Ruden, V. Bichko, B. Wahren, P. Pumpens, and L. Magnius.** 1994. Identification of hepatitis B virus core protein regions exposed or internalized at the surface of HBcAg particles by scanning with monoclonal antibodies. Virology **202:**912-20.

367. **Rasi, G., Mutchnick, M. G., Di Virgilio, D., Sinibaldi-Vallebona, P., Pierimarchi, P., Colella, F., Favalli, C. and Garcia, E.** 1996. Combination low dose lymphoblastoid interferon and thymosin alpha 1 therapy in the treatment of chronic hepatitis B. J Viral Hepat **3:**191-6.

368. **Ray, M. B., V. J. Desmet, A. F. Bradburne, J. Desmyter, J. Fevery, and J. De Groote.** 1976. Differential distribution of hepatitis B surface antigen and hepatitis B core antigen in the liver of hepatitis B patients. Gastroenterology **71:**462-9.

369. **Redd, A. J., N. Takezaki, S. T. Sherry, S. T. McGarvey, A. S. Sofro, and M. Stoneking.** 1995. Evolutionary history of the COII/tRNALys intergenic 9 base pair deletion in human mitochondrial DNAs from the Pacific. Mol Biol Evol **12:**604-15.

370. **Reesink, H. W., E. E. Reerink-Brongers, B. J. Lafeber-Schut, J. Kalshoven-Benschop, and H. G. Brummelhuis.** 1979. Prevention of chronic HBsAg carrier state in infants of HBsAg-positive mothers by hepatitis B immunoglobulin. Lancet **2:**436-8.

371. **Rehermann, B., C. Pasquinelli, S. M. Mosier, and F. V. Chisari.** 1995. Hepatitis B virus (HBV) sequence variation of cytotoxic T lymphocyte

epitopes is not common in patients with chronic HBV infection. J Clin Invest **96:**1527-34.

372. **Reignat, S., G. J. Webster, D. Brown, G. S. Ogg, A. King, S. L. Seneviratne, G. Dusheiko, R. Williams, M. K. Maini, and A. Bertoletti.** 2002. Escaping high viral load exhaustion: CD8 cells with altered tetramer binding in chronic hepatitis B virus infection. J Exp Med **195:**1089-101.

373. **Richardson, W. D., B. L. Roberts, and A. E. Smith.** 1986. Nuclear location signals in polyoma virus large-T. Cell **44:**77-85.

374. **Rigg, R. J. S., H.** 1992. Duck hepatitis B infection of hepatocytes is not dependent on low pH. J Virol **66:**2829-36.

375. **Rizzetto M, P. R., Gerin JL, Verme G.** 1997. Viral hepatitis and liver disease.

376. **Roberts, E. A.** 2000. Why treat chronic hepatitis B in childhood with interferon alpha? Gut **46:**591-3.

377. **Robertson, B. H.** 2001. Viral hepatitis and primates: historical and molecular analysis of human and nonhuman primate hepatitis A, B, and the GB-related viruses. J Viral Hepat **8:**233-42.

378. **Rodriguez-Frias, F., M. Buti, R. Jardi, M. Cotrina, L. Viladomiu, R. Esteban, and J. Guardia.** 1995. Hepatitis B virus infection: precore mutants and its relation to viral genotypes and core mutations. Hepatology **22:**1641-7.

379. **Roingeard, P., S. L. Lu, C. Sureau, M. Freschlin, B. Arbeille, M. Essex, and J. L. Romet-Lemonne.** 1990b. Immunocytochemical and electron microscopic study of hepatitis B virus antigen and complete particle production in hepatitis B virus DNA transfected HepG2 cells. Hepatology **11:**277-85.

380. **Roingeard, P., J. L. Romet-Lemonne, D. Leturcq, A. Goudeau, and M. Essex.** 1990a. Hepatitis B virus core antigen (HBc Ag) accumulation in an HBV nonproducer clone of HepG2-transfected cells is associated with cytopathic effect. Virology **179:**113-20.

381. **Rollier, C., C. Sunyach, L. Barraud, N. Madani, C. Jamard, C. Trepo, and L. Cova.** 1999. Protective and therapeutic effect of DNA-based immunization against hepadnavirus large envelope protein. Gastroenterology **116:**658-65.

382. **Roossinck, M. J., and A. Siddiqui.** 1987. In vivo phosphorylation and protein analysis of hepatitis B virus core antigen. J Virol **61:**955-61.

383. **Rosenberg, P. D. J. L.** 1999. Therapy with nucleoside analogues for hepatitis B infection. Clin Liver Dis **3:**349-361.

384. **Rosmorduc O, P. M., Capel F, Weber M, Brechot C, Kremsdorf D.** 1994. Expression of defective HBV DNA synthesis from a singly spliced HBV RNA leads to cytoplasmic accumulation of capsid and X proteins. Viral hepatitis and liver disease:92-96.

385. **Rossol, S., G. Marinos, P. Carucci, M. V. Singer, R. Williams, and N. V. Naoumov.** 1997. Interleukin-12 induction of Th1 cytokines is important for viral clearance in chronic hepatitis B. J Clin Invest **99:**3025-33.

386. **Ryu, C. J., D. Y. Cho, P. Gripon, H. S. Kim, C. Guguen-Guillouzo, and H. J. Hong.** 2000. An 80-kilodalton protein that binds to the pre-S1 domain of hepatitis B virus. J Virol **74:**110-6.

387. **Saito, T., T. Kamimura, M. Ishibashi, H. Shinzawa, and T. Takahashi.** 1992. Electron microscopic study of hepatitis B virus-associated antigens on the infected liver cell membrane in relation to analysis of immune target antigens in chronic hepatitis B. Gastroenterol Jpn **27:**734-44.

388. **Salfeld, J., E. Pfaff, M. Noah, and H. Schaller.** 1989. Antigenic determinants and functional domains in core antigen and e antigen from hepatitis B virus. J Virol **63:**798-808.

389. **Sallberg, M., P. Pushko, I. Berzinsh, V. Bichko, P. Sillekens, M. Noah, P. Pumpens, E. Grens, B. Wahren, and L. O. Magnius.** 1993. Immunochemical structure of the carboxy-terminal part of hepatitis B e antigen: identification of internal and surface-exposed sequences. J Gen Virol **74 (Pt 7):**1335-40.

390. **Sallberg, M., U. Ruden, B. Wahren, and L. O. Magnius.** 1994. Immune recognition of linear antigenic regions within the hepatitis B pre-C and C-gene translation products using synthetic peptides. J Med Virol **42:**7-15.

391. **Sallberg, M., U. Ruden, B. Wahren, M. Noah, and L. O. Magnius.** 1991. Human and murine B-cells recognize the HBeAg/beta (or HBe2) epitope as a linear determinant. Mol Immunol **28:**719-26.

392. **Sampliner, R. E., W. B. Bias, E. Carney, A. Hillis, and W. D. Hillis.** 1981. HLA antigens and HBV infection: evaluation in the chronic carrier state and in a large family. Tissue Antigens **18:**247-51.

393. **Samuel, C. E.** 2001. Antiviral actions of interferons. Clin Microbiol Rev **14:**778-809, table of contents.

394. **Samuel, D., A. Bismuth, D. Mathieu, J. L. Arulnaden, M. Reynes, J. P. Benhamou, C. Brechot, and H. Bismuth.** 1991. Passive immunoprophylaxis after liver transplantation in HBsAg-positive patients. Lancet **337:**813-5.

395. **Samuel, D., R. Muller, G. Alexander, L. Fassati, B. Ducot, J. P. Benhamou, and H. Bismuth.** 1993. Liver transplantation in European patients with the hepatitis B surface antigen. N Engl J Med **329:**1842-7.

396. **Sanchez-Tapias, J. M., J. Costa, A. Mas, M. Bruguera, and J. Rodes.** 2002. Influence of hepatitis B virus genotype on the long-term outcome of chronic hepatitis B in western patients. Gastroenterology **123:**1848-56.

397. **Santantonio, T., M. C. Jung, S. Miska, G. Pastore, G. R. Pape, and H. Will.** 1991a. Prevalence and type of pre-C HBV mutants in anti-HBe positive carriers with chronic liver disease in a highly endemic area. Virology **183:**840-4.

398. **Santantonio, T., M. C. Jung, R. Schneider, D. Fernholz, M. Milella, L. Monno, G. Pastore, G. R. Pape, and H. Will.** 1992. Hepatitis B virus genomes that cannot synthesize pre-S2 proteins occur frequently and as dominant virus populations in chronic carriers in Italy. Virology **188:**948-52.

399. **Santantonio, T., M. C. Jung, R. Schneider, G. Pastore, G. R. Pape, and H. Will.** 1991b. Selection for a pre-C stop codon mutation in a hepatitis B virus variant with a pre-C initiation codon mutation during interferon treatment. J Hepatol **13:**368-71.

400. **Scaglioni, P. P., M. Melegari, and J. R. Wands.** 1994. Characterization of hepatitis B virus core mutants that inhibit viral replication. Virology **205:**112-20.

401. **Schalm SW, H. J., Cianciara J, Farrell G, Sherman M, Willems B, Dhillon A, Moorat A, Barber J, Gray DF.** 2000. Lamivudine and alpha

interferon combination treatment of patients with chronic hepatitis B infection: a randomised trial. Gut **46:**562-8.

402. **Schlicht, H. J., R. Bartenschlager, and H. Schaller.** 1989. The duck hepatitis B virus core protein contains a highly phosphorylated C terminus that is essential for replication but not for RNA packaging. J Virol **63:**2995-3000.

403. **Schlicht, H. J., and G. Wasenauer.** 1991. The quaternary structure, antigenicity, and aggregational behavior of the secretory core protein of human hepatitis B virus are determined by its signal sequence. J Virol **65:**6817-25.

404. **Schodel, F., A. M. Moriarty, D. L. Peterson, J. A. Zheng, J. L. Hughes, H. Will, D. J. Leturcq, J. S. McGee, and D. R. Milich.** 1992. The position of heterologous epitopes inserted in hepatitis B virus core particles determines their immunogenicity. J Virol **66:**106-14.

405. **Seeger, C., B. Baldwin, and B. C. Tennant.** 1989. Expression of infectious woodchuck hepatitis virus in murine and avian fibroblasts. J Virol **63:**4665-9.

406. **Seeger, C., and W. S. Mason.** 2000. Hepatitis B virus biology. Microbiol Mol Biol Rev **64:**51-68.

407. **Seifer, M., S. Zhou, and D. N. Standring.** 1993. A micromolar pool of antigenically distinct precursors is required to initiate cooperative assembly of hepatitis B virus capsids in Xenopus oocytes. J Virol **67:**249-57.

408. **Serinoz, E., M. Varli, E. Erden, K. Cinar, A. Kansu, O. Uzunalimoglu, C. Yurdaydin, and H. Bozkaya.** 2003. Nuclear localization of hepatitis B core antigen and its relations to liver injury, hepatocyte proliferation, and viral load. J Clin Gastroenterol **36:**269-72.

409. **Sewell, A. K., D. A. Price, A. Oxenius, A. D. Kelleher, and R. E. Phillips.** 2000. Cytotoxic T lymphocyte responses to human immunodeficiency virus: control and escape. Stem Cells **18:**230-44.

410. **Sharma, R. R., R. K. Dhiman, Y. Chawla, and R. K. Vasistha.** 2002. Immunohistochemistry for core and surface antigens in chronic hepatitis. Trop Gastroenterol **23:**16-9.

411. **Shih, C. H., L. S. Li, S. Roychoudhury, and M. H. Ho.** 1989. In vitro propagation of human hepatitis B virus in a rat hepatoma cell line. Proc Natl Acad Sci U S A **86:**6323-7.

412. **Shiina, S., H. Fujino, T. Kawabe, K. Tagawa, T. Unuma, M. Yoneyama, T. Ohmori, S. Suzuki, M. Kurita, and Y. Ohashi.** 1991. Relationship of HBsAg subtypes with HBeAg/anti-HBe status and chronic liver disease. Part II: Evaluation of epidemiological factors and suspected risk factors of liver dysfunction. Am J Gastroenterol **86:**872-5.

413. **Shimada, N., K. Yamamoto, M. J. Kuroda, R. Terada, T. Hakoda, H. Shimomura, H. Hata, E. Nakayama, and Y. Shiratori.** 2003. HBcAg-specific CD8 T cells play an important role in virus suppression, and acute flare-up is associated with the expansion of activated memory T cells. J Clin Immunol **23:**223-32.

414. **Siddiqui, A., S. Jameel, and J. Mapoles.** 1986. Transcriptional control elements of hepatitis B surface antigen gene. Proc Natl Acad Sci U S A **83:**566-70.

415. **Simmonds, P.** 2001b. The origin and evolution of hepatitis viruses in humans. J Gen Virol **82:**693-712.

416. **Simmonds, P.** 2001a. Reconstructing the origins of human hepatitis viruses. Philos Trans R Soc Lond B Biol Sci **356:**1013-26.

417. **Sobao, Y., K. Sugi, H. Tomiyama, S. Saito, S. Fujiyama, M. Morimoto, S. Hasuike, H. Tsubouchi, K. Tanaka, and M. Takiguch.** 2001. Identification of hepatitis B virus-specific CTL epitopes presented by HLA-A*2402, the most common HLA class I allele in East Asia. J Hepatol **34:**922-9.

418. **Sohn, S., I. Jaitovitch-Groisman, N. Benlimame, J. Galipeau, G. Batist, and M. A. Alaoui-Jamali.** 2000. Retroviral expression of the hepatitis B virus x gene promotes liver cell susceptibility to carcinogen-induced site specific mutagenesis. Mutat Res **460:**17-28.

419. **Staprans, S., D. D. Loeb, and D. Ganem.** 1991. Mutations affecting hepadnavirus plus-strand DNA synthesis dissociate primer cleavage from translocation and reveal the origin of linear viral DNA. J Virol **65:**1255-62.

420. **Starkman, S. E., D. M. MacDonald, J. C. Lewis, E. C. Holmes, and P. Simmonds.** 2003. Geographic and species association of hepatitis B virus genotypes in non-human primates. Virology **314:**381-93.

421. **Sterneck, M., S. Gunther, J. Gerlach, N. V. Naoumov, T. Santantonio, L. Fischer, X. Rogiers, H. Greten, R. Williams, and H. Will.**

1997. Hepatitis B virus sequence changes evolving in liver transplant recipients with fulminant hepatitis. J Hepatol **26:**754-64.

422. **Sterneck, M., T. Kalinina, S. Otto, S. Gunther, L. Fischer, M. Burdelski, H. Greten, C. E. Broelsch, and H. Will.** 1998. Neonatal fulminant hepatitis B: structural and functional analysis of complete hepatitis B virus genomes from mother and infant. J Infect Dis **177:**1378-81.

423. **Stuyver, L., S. De Gendt, C. Van Geyt, F. Zoulim, M. Fried, R. F. Schinazi, and R. Rossau.** 2000. A new genotype of hepatitis B virus: complete genome and phylogenetic relatedness. J Gen Virol **81:**67-74.

424. **Su, Q., C. H. Schroder, W. J. Hofmann, G. Otto, R. Pichlmayr, and P. Bannasch.** 1998. Expression of hepatitis B virus X protein in HBV-infected human livers and hepatocellular carcinomas. Hepatology **27:**1109-20.

425. **Sugauchi, F., E. Orito, T. Ichida, H. Kato, H. Sakugawa, S. Kakumu, T. Ishida, A. Chutaputti, C. L. Lai, R. G. Gish, R. Ueda, Y. Miyakawa, and M. Mizokami.** 2003. Epidemiologic and virologic characteristics of hepatitis B virus genotype B having the recombination with genotype C. Gastroenterology **124:**925-32.

426. **Sugauchi, F., E. Orito, T. Ichida, H. Kato, H. Sakugawa, S. Kakumu, T. Ishida, A. Chutaputti, C. L. Lai, R. Ueda, Y. Miyakawa, and M. Mizokami.** 2002. Hepatitis B virus of genotype B with or without recombination with genotype C over the precore region plus the core gene. J Virol **76:**5985-92.

427. **Summers, J., and W. S. Mason.** 1982. Replication of the genome of a hepatitis B--like virus by reverse transcription of an RNA intermediate. Cell **29:**403-15.

428. **Summers, J., P. M. Smith, and A. L. Horwich.** 1990. Hepadnavirus envelope proteins regulate covalently closed circular DNA amplification. J Virol **64:**2819-24.

429. **Suri, D., R. Schilling, A. R. Lopes, I. Mullerova, G. Colucci, R. Williams, and N. V. Naoumov.** 2001. Non-cytolytic inhibition of hepatitis B virus replication in human hepatocytes. J Hepatol **35:**790-7.

430. **Suzuki, F., Y. Suzuki, A. Tsubota, N. Akuta, T. Someya, M. Kobayashi, S. Saitoh, Y. Arase, K. Ikeda, and H. Kumada.** 2002. Mutations

of polymerase, precore and core promoter gene in hepatitis B virus during 5-year lamivudine therapy. J Hepatol **37:**824-30.

431. **Suzuki, Y., H. Kumada, K. Ikeda, K. Chayama, Y. Arase, S. Saitoh, A. Tsubota, M. Kobayashi, M. Koike, N. Ogawa, and K. Tanikawa.** 1999. Histological changes in liver biopsies after one year of lamivudine treatment in patients with chronic hepatitis B infection. J Hepatol **30:**743-8.

432. **Takada, S., N. Kaneniwa, N. Tsuchida, and K. Koike.** 1996. Hepatitis B virus X gene expression is activated by X protein but repressed by p53 tumor suppressor gene product in the transient expression system. Virology **216:**80-9.

433. **Takahashi, K., Y. Akahane, K. Hino, Y. Ohta, and S. Mishiro.** 1998. Hepatitis B virus genomic sequence in the circulation of hepatocellular carcinoma patients: comparative analysis of 40 full-length isolates. Arch Virol **143:**2313-26.

434. **Takahashi, K., B. Brotman, S. Usuda, S. Mishiro, and A. M. Prince.** 2000. Full-genome sequence analyses of hepatitis B virus (HBV) strains recovered from chimpanzees infected in the wild: implications for an origin of HBV. Virology **267:**58-64.

435. **Takahashi, K., A. Machida, G. Funatsu, M. Nomura, S. Usuda, S. Aoyagi, K. Tachibana, H. Miyamoto, M. Imai, T. Nakamura, Y. Miyakawa, and M. Mayumi.** 1983. Immunochemical structure of hepatitis B e antigen in the serum. J Immunol **130:**2903-7.

436. **Telenta, P. F., G. P. Poggio, J. L. Lopez, J. Gonzalez, A. Lemberg, and R. H. Campos.** 1997. Increased prevalence of genotype F hepatitis B virus isolates in Buenos Aires, Argentina. J Clin Microbiol **35:**1873-5.

437. **ter Borg, F., F. J. ten Kate, H. T. Cuypers, A. Leentvaar-Kuijpers, J. Oosting, P. M. Wertheim-van Dillen, P. Honkoop, M. C. Rasch, R. A. de Man, J. van Hattum, R. A. Chamuleau, G. N. Tytgat, and E. A. Jones.** 2000. A survey of liver pathology in needle biopsies from HBsAg and anti-HBe positive individuals. J Clin Pathol **53:**541-8.

438. **Thimme, R., S. Wieland, C. Steiger, J. Ghrayeb, K. A. Reimann, R. H. Purcell, and F. V. Chisari.** 2003. CD8(+) T cells mediate viral clearance and disease pathogenesis during acute hepatitis B virus infection. J Virol **77:**68-76.

439. **Thio, C. L., M. Carrington, D. Marti, S. J. O'Brien, D. Vlahov, K. E. Nelson, J. Astemborski, and D. L. Thomas.** 1999. Class II HLA alleles and hepatitis B virus persistence in African Americans. J Infect Dis **179:**1004-6.

440. **Thursz, M. R., D. Kwiatkowski, C. E. Allsopp, B. M. Greenwood, H. C. Thomas, and A. V. Hill.** 1995. Association between an MHC class II allele and clearance of hepatitis B virus in the Gambia. N Engl J Med **332:**1065-9.

441. **Tipples, G. A., M. M. Ma, K. P. Fischer, V. G. Bain, N. M. Kneteman, and D. L. Tyrrell.** 1996. Mutation in HBV RNA-dependent DNA polymerase confers resistance to lamivudine in vivo. Hepatology **24:**714-7.

442. **Tong, S., J. Li, and J. R. Wands.** 1999. Carboxypeptidase D is an avian hepatitis B virus receptor. J Virol **73:**8696-702.

443. **Tong, S. P., J. S. Li, L. Vitvitski, and C. Trepo.** 1990. Active hepatitis B virus replication in the presence of anti-HBe is associated with viral variants containing an inactive pre-C region. Virology **176:**596-603.

444. **Tordjeman, M., G. Fontan, V. Rabillon, J. Martin, C. Trepo, A. Hoffenbach, K. Mabrouk, J. M. Sabatier, J. Van Rietschoten, and G. Somme.** 1993. Characterization of minor and major antigenic regions within the hepatitis B virus nucleocapsid. J Med Virol **41:**221-9.

445. **Torresi, J.** 2002. The virological and clinical significance of mutations in the overlapping envelope and polymerase genes of hepatitis B virus. J Clin Virol **25:**97-106.

446. **Treichel, U., K. H. Meyer zum Buschenfelde, H. P. Dienes, and G. Gerken.** 1997. Receptor-mediated entry of hepatitis B virus particles into liver cells. Arch Virol **142:**493-8.

447. **Treinin, M., and O. Laub.** 1987. Identification of a promoter element located upstream from the hepatitis B virus X gene. Mol Cell Biol **7:**545-8.

448. **Tsubota, A., H. Kumada, K. Takaki, K. Chayama, M. Kobayashi, Y. Suzuki, S. Saitoh, Y. Arase, N. Murashima, and K. Ikeda.** 1998. Deletions in the hepatitis B virus core gene may influence the clinical outcome in hepatitis B e antigen-positive asymptomatic healthy carriers. J Med Virol **56:**287-93.

449. **Tsui, L. V., L. G. Guidotti, T. Ishikawa, and F. V. Chisari.** 1995. Posttranscriptional clearance of hepatitis B virus RNA by cytotoxic T lymphocyte-activated hepatocytes. Proc Natl Acad Sci U S A **92:**12398-402.

450. **Tsurimoto, T., A. Fujiyama, and K. Matsubara.** 1987. Stable expression and replication of hepatitis B virus genome in an integrated state in a human hepatoma cell line transfected with the cloned viral DNA. Proc Natl Acad Sci U S A **84:**444-8.

451. **Tuttleman, J. S., C. Pourcel, and J. Summers.** 1986. Formation of the pool of covalently closed circular viral DNA in hepadnavirus-infected cells. Cell **47:**451-60.

452. **Underwood, M. R., and H. M. Fried.** 1990. Characterization of nuclear localizing sequences derived from yeast ribosomal protein L29. Embo J **9:**91-9.

453. **Uy, A., V. Bruss, W. H. Gerlich, H. G. Kochel, and R. Thomssen.** 1986. Precore sequence of hepatitis B virus inducing e antigen and membrane association of the viral core protein. Virology **155:**89-96.

454. **Vail, B. A.** 1997. Management of chronic viral hepatitis. Am Fam Physician **55:**2749-56, 2759-61.

455. **van den Oord, J. J., R. de Vos, and V. J. Desmet.** 1986. In situ distribution of major histocompatibility complex products and viral antigens in chronic hepatitis B virus infection: evidence that HBc-containing hepatocytes may express HLA-DR antigens. Hepatology **6:**981-9.

456. **Vilcek J, S. G.** 1996. Interferons and other cytokines, vol. 1. Lippincott-Raven, Philadelphia.

457. **von Weizsacker, F., H. E. Blum, and J. R. Wands.** 1992. Cleavage of hepatitis B virus RNA by three ribozymes transcribed from a single DNA template. Biochem Biophys Res Commun **189:**743-8.

458. **Wai, C. T., C. J. Chu, M. Hussain, and A. S. Lok.** 2002. HBV genotype B is associated with better response to interferon therapy in HBeAg(+) chronic hepatitis than genotype C. Hepatology **36:**1425-30.

459. **Wakita, T., S. Kakumu, M. Shibata, K. Yoshioka, Y. Ito, T. Shinagawa, T. Ishikawa, M. Takayanagi, and T. Morishima.** 1991. Detection of pre-C and core region mutants of hepatitis B virus in chronic hepatitis B virus carriers. J Clin Invest **88:**1793-801.

460. **Walker, M. P., and W. I. Lipkin.** 2002. Characterization of the nuclear localization signal of the borna disease virus polymerase. J Virol **76:**8460-7.

461. **Wallace, L. A., Carman WF.** 1997. Surface gene variation of HBV: Scientific and medical relevance. Viral hepatitis **3:**5-16.

462. **Wallace, L. E., D. J. Moss, A. B. Rickinson, A. J. McMichael, and M. A. Epstein.** 1981. Cytotoxic T cell recognition of Epstein-Barr virus-infected B cells. II. Blocking studies with monoclonal antibodies to HLA determinants. Eur J Immunol **11:**694-9.

463. **Walsh, K. M., T. Woodall, P. Lamy, D. G. Wight, S. Bloor, and G. J. Alexander.** 2001. Successful treatment with adefovir dipivoxil in a patient with fibrosing cholestatic hepatitis and lamivudine resistant hepatitis B virus. Gut **49:**436-40.

464. **Wang, G. H., and C. Seeger.** 1993. Novel mechanism for reverse transcription in hepatitis B viruses. J Virol **67:**6507-12.

465. **Wang, W. L., W. T. London, and M. A. Feitelson.** 1991. Hepatitis B x antigen in hepatitis B virus carrier patients with liver cancer. Cancer Res **51:**4971-7.

466. **Warren, K. S., J. L. Heeney, R. A. Swan, Heriyanto, and E. J. Verschoor.** 1999. A new group of hepadnaviruses naturally infecting orangutans (Pongo pygmaeus). J Virol **73:**7860-5.

467. **Waters, J. A., T. P. Jowett, and H. C. Thomas.** 1986. Identification of a dominant immunogenic epitope of the nucleocapsid (HBc) of the hepatitis B virus. J Med Virol **19:**79-86.

468. **Welch, P. J., R. Tritz, S. Yei, J. Barber, and M. Yu.** 1997. Intracellular application of hairpin ribozyme genes against hepatitis B virus. Gene Ther **4:**736-43.

469. **Wilson, J. N., D. J. Nokes, and W. F. Carman.** 1999. The predicted pattern of emergence of vaccine-resistant hepatitis B: a cause for concern? Vaccine **17:**973-8.

470. **Wingfield, P. T., S. J. Stahl, R. W. Williams, and A. C. Steven.** 1995. Hepatitis core antigen produced in Escherichia coli: subunit composition, conformational analysis, and in vitro capsid assembly. Biochemistry **34:**4919-32.

471. **Wright, T. L. T., M. J. Hsu, H. H.** 1999. Phase 1 study of a potent adjuanted hepatitis B vaccine (HBV/MF59) for therapy of chronic hepatitis. Hepatology **30:**421A.

472. **Wu, J. C., S. D. Lee, J. Y. Wang, L. P. Ting, Y. T. Tsai, S. H. Tsay, and M. J. Tong.** 1987. Correlation between hepatic hepatitis B core antigen and serum hepatitis B virus-DNA levels in patients with chronic hepatitis B virus infections in Taiwan. Arch Pathol Lab Med **111:**181-4.

473. **Wu, P. C., J. W. Fang, C. L. Lai, S. S. Lau, C. K. Lo, A. Lai, and J. Y. Lau.** 1996. Hepatic expression of hepatitis B virus genome in chronic hepatitis B virus infection. Am J Clin Pathol **105:**87-95.

474. **Wu, P. C., J. Y. Lau, T. K. Lau, S. K. Lau, and C. L. Lai.** 1993. Relationship between intrahepatic expression of hepatitis B viral antigens and histology in Chinese patients with chronic hepatitis B virus infection. Am J Clin Pathol **100:**648-53.

475. **Wynne, S. A., R. A. Crowther, and A. G. Leslie.** 1999. The crystal structure of the human hepatitis B virus capsid. Mol Cell **3:**771-80.

476. **Xu ZY, L. C., Francis DP, Purcell RH, Gun ZL, Duan SC, Chen RJ, Margolis HS, Huang CH, Maynard JE.** 1985. Prevention of perinatal acquisition of hepatitis B virus carriage using vaccine: preliminary report of a randomized, double-blind placebo-controlled and comparative trial. Pediatrics **76:**713-18.

477. **Yaginuma, K., Y. Shirakata, M. Kobayashi, and K. Koike.** 1987. Hepatitis B virus (HBV) particles are produced in a cell culture system by transient expression of transfected HBV DNA. Proc Natl Acad Sci U S A **84:**2678-82.

478. **Yamada, G., and P. K. Nakane.** 1977. Hepatitis B core and surface antigens in liver tissue. Light and electron microscopic localization by the peroxidase-labeled antibody method. Lab Invest **36:**649-59.

479. **Yamauchi, K., T. Nakamura, H. Yonemitsu, H. Sekiya, J. Katoh, and H. Obata.** 1993. Possible role of preS2 peptides presented by MHC class I antigen in the pathogenesis of chronic hepatitis B. J Hepatol **17:**S6-9.

480. **Yamishita, Y., S. Kurashina, Y. Miyakawa, and M. Mayumi.** 1975. South-to-north gradient in distribution of the r determinant of hepatitis B surface antigen in Japan. J Infect Dis **131:**567-9.

481. **Yang, Z., I. J. Lauder, and H. J. Lin.** 1995. Molecular evolution of the hepatitis B virus genome. J Mol Evol **41:**587-96.

482. **Yeh, C. T., H. T. Chiu, C. M. Chu, and Y. F. Liaw.** 1998. G1 phase dependent nuclear localization of relaxed-circular hepatitis B virus DNA and aphidicolin-induced accumulation of covalently closed circular DNA. J Med Virol **55:**42-50.

483. **Yeh, C. T., C. M. Chu, and Y. F. Liaw.** 1995. A single serine mutation on the nuclear localization signal of hepatitis B virus core protein abolishes the inhibition of nuclear transport by surface proteins. Biochem Biophys Res Commun **213:**1068-74.

484. **Yeh, C. T., L. H. Hong, J. H. Ou, C. M. Chu, and Y. F. Liaw.** 1996. Characterization of nuclear localization of a hepatitis B virus precore protein derivative P22. Arch Virol **141:**425-38.

485. **Yeh, C. T., Y. F. Liaw, and J. H. Ou.** 1990. The arginine-rich domain of hepatitis B virus precore and core proteins contains a signal for nuclear transport. J Virol **64:**6141-7.

486. **Yeh, C. T., and J. H. Ou.** 1991. Phosphorylation of hepatitis B virus precore and core proteins. J Virol **65:**2327-31.

487. **Yeh, C. T., J. H. Ou, C. M. Chu, and Y. F. Liaw.** 1994. Alteration of the subcellular localization of hepatitis B virus core protein by large but not small surface proteins. Biochem Biophys Res Commun **203:**1348-54.

488. **Yeh, C. T., S. W. Wong, Y. K. Fung, and J. H. Ou.** 1993. Cell cycle regulation of nuclear localization of hepatitis B virus core protein. Proc Natl Acad Sci U S A **90:**6459-63.

489. **Yoo, B. C., J. W. Park, H. J. Kim, D. H. Lee, Y. J. Cha, and S. M. Park.** 2003. Precore and core promoter mutations of hepatitis B virus and hepatitis B e antigen-negative chronic hepatitis B in Korea. J Hepatol **38:**98-103.

490. **Yu, M., and J. Summers.** 1994. Multiple functions of capsid protein phosphorylation in duck hepatitis B virus replication. J Virol **68:**4341-8.

491. **Yuan, T. T., and C. Shih.** 2000. A frequent, naturally occurring mutation (P130T) of human hepatitis B virus core antigen is compensatory for immature secretion phenotype of another frequent variant (I97L). J Virol **74:**4929-32.

492. **Yuan, T. T., P. C. Tai, and C. Shih.** 1999. Subtype-independent immature secretion and subtype-dependent replication deficiency of a highly

frequent, naturally occurring mutation of human hepatitis B virus core antigen. J Virol **73:**10122-8.

493. **Yuen, M. F., E. Sablon, D. K. Wong, H. J. Yuan, B. C. Wong, A. O. Chan, and C. L. Lai.** 2003. Role of hepatitis B virus genotypes in chronic hepatitis B exacerbation. Clin Infect Dis **37:**593-7.

494. **Zampino, R., S. Lobello, M. Chiaramonte, C. Venturi-Pasini, U. Dumpis, M. Thursz, and P. Karayiannis.** 2002. Intra-familial transmission of hepatitis B virus in Italy: phylogenetic sequence analysis and amino-acid variation of the core gene. J Hepatol **36:**248-53.

495. **Zhang, H. G., T. Zhou, P. Yang, C. K. Edwards, 3rd, D. T. Curiel, and J. D. Mountz.** 1998. Inhibition of tumor necrosis factor alpha decreases inflammation and prolongs adenovirus gene expression in lung and liver. Hum Gene Ther **9:**1875-84.

496. **Zhang, H. G., Zhou T, Yang P, Edwards CK, Curiel DT, Mountz JD.** 1998. Inhibition of TNF alpha decrease inflammation and prolongs adenovirus gene expression in lung and liver. Hum Gene Ther **9:**1875-1884.

497. **Zhong, S., J. Y. Chan, W. Yeo, J. S. Tam, and P. J. Johnson.** 2000. Frequent integration of precore/core mutants of hepatitis B virus in human hepatocellular carcinoma tissues. J Viral Hepat **7:**115-23.

498. **Zhu, M., W. T. London, L. X. Duan, and M. A. Feitelson.** 1993. The value of hepatitis B x antigen as a prognostic marker in the development of hepatocellular carcinoma. Int J Cancer **55:**571-6.

499. **Zlotnick, A., N. Cheng, S. J. Stahl, J. F. Conway, A. C. Steven, and P. T. Wingfield.** 1997. Localization of the C terminus of the assembly domain of hepatitis B virus capsid protein: implications for morphogenesis and organization of encapsidated RNA. Proc Natl Acad Sci U S A **94:**9556-61.

500. **Zollner, B., J. Petersen, P. Schafer, M. Schroter, R. Laufs, M. Sterneck, and H. H. Feucht.** 2002. Subtype-dependent response of hepatitis B virus during the early phase of lamivudine treatment. Clin Infect Dis **34:**1273-7.

501. **Zollner, B., J. Petersen, M. Schroter, R. Laufs, V. Schoder, and H. H. Feucht.** 2001. 20-fold increase in risk of lamivudine resistance in hepatitis B virus subtype adw. Lancet **357:**934-5.

502. **Zvibel, I., A. S. Fiorino, S. Brill, and L. M. Reid.** 1998. Phenotypic characterization of rat hepatoma cell lines and lineage-specific regulation of gene expression by differentiation agents. Differentiation **63:**215-23.

Journal of Viral Hepatitis, 2004, **11**, 1–14

HBV core sequence: definition of genotype-specific variability and correlation with geographical origin

M. Jazayeri,[1] A. A. Basuni,[1] N. Sran,[2] R. Gish,[2] G. Cooksley,[3] S. Locarnini[4] and W. F. Carman[1,5] [1]*Division of Virology, Institute of Biomedical and Life Science, University of Glasgow, Glasgow, UK;* [2]*Departments of Medicine and Transplantation, California Pacific Medical Center, Sanfrancisco, CA, USA;* [3]*Clinical Research Center, Royal Brisbane Hospital Foundation, Brisbane, Australia;* [4]*Research and Molecular Development, Victorian Infectious Diseases Reference Laboratory, Melbourne, Australia; and* [5]*West of Scotland specialist Virology Center, Gartnavel General Hospital, Glasgow, UK*

Received xxxx 200x; accepted for publication xxxx 200x

SUMMARY. There are eight genotypes and nine subtypes of HBV. Small differences in geographical origin are associated with sequence changes in the surface gene. Here, we compared core gene sequences from different genotypes and geographical regions. Specific combinations of 24 amino acid substitutions at nine residues allowed allocation of a sequence to a subtype. Six of these nine residues were located in different T cell epitopes depending on HBV geographical area and/or genotype. Thirty-four nucleotide changes were associated uniquely with specific genotypes and subtypes. Unique amino acid and nucleotide variants were found in a majority of sequences from specific countries as well as within subtype ayw2 and adr. Specific nucleotide motifs were defined for Korean, Indian, Chinese, Italian and Pacific region isolates. Finally, we observed amino acid motifs that were common to either South-east Asian or Western populations, irrespective of subtype. We believe that HBV strains spread within constrained ethnic groups, result in selection pressures that define sequence variability within each subtype. It suggests that particular T cell epitopes are specific for geographical regions, and thus ethnic groups; this may affect the design of immunomodulatory therapies.

Keywords: C gene variability, HBcAg, HBV genotypes, HBV subtypes.

INTRODUCTION

Hepatitis B virus (HBV) DNA contains four open reading frames. The core (C)-gene encodes the 183–185 amino acid long nucleocapsid protein and is preceded by the precore region [1]. Hepatitis B core antigen (HBcAg) contains helper T cell and humoural epitopes [2], and is an immunologic target of cytotoxic T lymphocytes (CTL) [1,3]. Different HLA-restricted HBcAg T cell epitopes have been proposed throughout HBcAg [4–8], which might be different on account of the diverse distribution of HLA antigens in different countries [5,9–10].

The HBV can be classified into at least eight genotypes (A–H) [11–14]. One common antigenic determinant, a, and two pairs of mutually exclusive determinants, d/y and w/r, initially identified four HBsAg subtypes. Later research defined nine subtypes in total: ayw1, ayw2, ayw3, ayw4, ayr, adw2, adw3, adw4, adrq+ and adrq– [14,15]. The relationship of these nine subtypes to genotypes A–G has been established and some subtypes can be found in more than one genotype. HBV genotypes have a characteristic geographical distribution, largely in agreement with subtype distribution [16].

Although the amino acid sequence of the core region is relatively conserved (compared with surface gene), amino acid substitutions have been observed in several reports [1–3,5,17–21]. This is largely linked to the clinical or serological picture [1–3,5,7,20]. However, in view of the heterogeneity of HBV strains, it can be difficult to determine whether substitutions recorded only once represent naturally occurring variants or merely an immune-selected mutation in that patient. So far, no systematic study has been conducted to correlate HBcAg diversity with HBV subtypes and ethnic background. This background information would not only be critical to the correct interpretation of sequences observed in clinical studies, but may aid in the design of immunotherapies specific for persons of a particular ethnic origin.

Here, HBV DNA in 91 sera from nine diverse geographical origins was analysed by sequencing the core gene.

Abbreviations:CTL, cytotoxic T lymphocyte; HBcAg, hepatitis B core antigen; HBV, hepatitis B virus; HLA, human leucocyte antigen.

Correspondence: William F. Carman, Gartnavel General Hospital, PO Box 16766, Glasgow G12 0ZA, Scotland. Tel.: +44 (0) 141-2110085; E-mail: w.carman@vir.gla.ac.uk

J V H | 5 3 4 | B | Dispatch: 12.4.04 | Journal: JVH | CE: Benjamin
Journal Name | Manuscript No. | Author Received: | No. of pages: 14 | PE: Raymond

Subtyping by additional S gene sequencing allowed definition of subtype-specific variants. Our sequences were then compared with those in international databases to add strength to the conclusions. We also found that nucleotide and/or amino acid variation can be correlated with geographical/ethnic background.

METHODS

Sera

Sera from 91 HBsAg positive patients from nine countries, including 49 samples from California (39 from Asian origin living in California and 10 from Caucasians), 12 from Italy, and 10 each from Scotland, India, and the Pacific region were selected (Table 1). All were positive for HBsAg with a variety of clinical features and serological status for HBeAg/anti-HBe. The Asian origin samples from California were previously sequenced for a precore variability study (results not shown). These findings then were compared with sequences of known geographical origin in the database to strengthen the associations.

Table 1 Origin of 91 HBsAg-positive sera that were used as the source for HBcAg sequencing

Country	No. of samples
South-east Asia	
China	20
Korea	8
Japan	5
Vietnam	6
Pacific	10
Caucasian	
USA	10
Scotland	10
Italy	12
India	10
Total	91

DNA extraction

A 50 μL aliquot of serum was mixed with 150 μL of PBS, and then 200 μL binding buffer (High Pure Viral Nucleic Acid Kit, Roche, Germany) supplemented with poly A carrier RNA was added. The sample was digested with 40 μL proteinase K and incubated at 72 °C for 10 min. After adding 100 μL of isopropanol, the mixture was applied to a filter tube containing fleece and washed twice with wash buffer. After centrifugation, 50 μL of elution buffer was added and the eluted DNA was stored at −20 °C.

Polymerase chain reaction

Polymerase chain reaction (PCR) reactions were carried out in 50 μL of a mixture containing 5 μL of the extracted DNA, using standard methodology. The complete core gene was amplified using C1, C3A and C4 primers (Table 2). In addition, the region of surface gene specifying HBV subtype (amino acid positions 122–160) was amplified using S1, S2Na, S6C and S7D primers (Table 2). First, round PCR was performed using 1 U of *Taq* DNA polymerase (Life Technologies, Paisley, UK), 1.4 of *Taq* start antibody (Clontech Laboratories Inc., Palo Alto, CA, USA), 0.25 mM of each dNTP (Pharmacia, St Albans, UK), 10x reaction buffer, 25 pmol of C1 and C4 primers (for core gene) and 12.5 pmol of S1 and S2 primers (for surface gene). For the second round PCR, 1 μL of first round PCR product was added to 49 μL of the reaction mixture with the same composition as the first round except that C1 was replaced by C3A, and S1/S2Na, were replaced by S6C/S7D. A quantity of 5 μL of the second round PCR products were analysed by electrophoresis in 1% agarose gel, stained by ethidium bromide, and visualized under u.v. light.

The PCR products of the correct size were purified using the High Pure PCR Product Purification Kit (Roche Diagnostics, Penzberg, Germany), following the manufacturer's instruction.

Primer	Gene	Sequence 5′ to 3′ of oligonucleotide	Base position
C1	Core	CGG GAT CCG AGG AGT TGG GGG AGG AGA TT	1726–1754
C3A	Core	G(AG)T CTR TGT AWT AGG AGG CTG	1763–1783
C4	Core	CCT TAT GAG TCC AAG G(AG)A TA	2478–2459
S1	Surface	CCT GCT GGT GGC TCC AGT TC	56–75
S2Na	Surface	CCA CAA TTC (K)TT GAC ATA CTT TCC A	1003–979
S6C	Surface	GCA CAC GGA ATT CCG AGG ACT GGG GAC CCT G	113–146
S7D	Surface	GAC ACC AAG CTT GGT TAG GGT TTA AAT GTA TAC C	857–823

Table 2 Oligonucleotide primers used for PCR and sequencing. Base positions numbered from the *Eco*RI site. W and K represent mixed base sites, A or T and G or T respectively

DNA sequencing

The HBsAg subtype of the sequences was defined by substitutions in the 'a' determinant between codons 122 and 160 inclusive. Direct sequencing of core and surface genes was carried out (Perkin Elmer ABI-PRISM™ 377 DNA Sequencer, Fostercity, CA, USA) using 2 pmol of appropriate primers: C3A and C4 for the core gene; S6C and S7D for surface gene. Because of unavailability of DNA materials, we did not sequence the S gene of Indian samples. The results were analysed using Sequence Navigator software.

Sequence analysis

After allocating a sequence to an HBV genotype by analysis of the S gene, the core gene amino acid/nucleotide variations that were found uniquely in that subtype or genotype were recorded. Variants found in a majority of HBcAg sequences (>50%) from a particular geographical area were confirmed with HBcAg sequences from the same area obtained from EMBL, GenBank and NCBI (Table 3). Sequences were only selected from databases if the subtype, country of origin and serological picture were provided. Finally, geography-specific nucleotide and amino acid variations in each genotype were identified. To qualify, they had to be both unique to a geographical region and found in a majority of sequences from that region. Sequences have been submitted to GenBank, numbered from AF323463 to AF323470 and from AF324066 to AF324148.

Phylogenetic analysis

Sequences of core and surface genes were aligned using the BioEdit Package version 5.0.9 [22], and a neighbour-joining phylogenetic tree constructed using the Treecon Package [23] employing a Kimura distance matrix [24]. Associations were tested by bootstrap re-sampling analysis using 100 replicates [25]. Associations with a bootstrap value of greater than 70% were deemed significant.

RESULTS

Phylogenetic analysis

Figure 1a,b shows that sequences grouped into six major clusters. Four of these were occupied by genotypes A–D. Pacific sequences composed two additional clusters (in both C and S trees), one in genotypes C and one in D. Compared with the rest, the most homogeneous groups was the Pacific sequences. In genotype D strains, most of the Indian sequences comprised a subcluster in the C tree. Similarly, a majority of the strains found in Scotland and the USA (genotype A), were homogenous in both C and S trees. The phylogenetic trees constructed based on S and C genes revealed that the topological features of all strains except of A4, A5, I54 and I113 in the phylogenetic trees were identical. The C gene of A5 grouped with genotype B sequences, whereas the S gene with genotype C ones (the Pacific adrq−). Similarly, strains I54 and I113, based on core sequencing, grouped with genotype D, whereas, the S genes with genotype A.

Genotype/subtype-specific nucleotide substitution

At the nucleotide level, it was possible to identify unique variants for subtypes adw2 (belonging to genotypes A and B), ayw1, ayw2, adr and adrq− (there was no ayr subtype in our samples). Overall, 34 genotype-specific nucleotide substitutions in 20 positions were found (Table 4), of which 18 were silent (synonymous) and 14 were missense (nonsynonymous). Of 34 variations, it was possible to define nine unique nucleotide variants for specific subtypes: C_{2121} and G_{2242} for ayw2; C_{2242} for ayw3 strains; T_{1966}, G_{1975} and G_{2075} for adrq−; and C_{2189} for adrq+. A_{2354} was unique for adw2 and A_{2304} was unique for ayw1, both belonging to genotype B. Our results concurred with those of sequences in databases and we assume that some of the amino acid and/or nucleotide differences between our results and other reports could be related to the diversity in severity of chronic disease in specific patients.

Amino acid substitutions and genotype/subtype

The amino acid residues at nine codons (aa 27, 59, 67, 74, 83, 87, 91, 97 and 116), and an insertion at amino acid position 152, allowed allocation of a specific subtype (Table 5). Of these nine residues, the following were unique for genotypes and their corresponding subtypes: N74 and N87 in genotype A (adw2); F97 and I116 in genotype D; V59 and I91 in genotype C strains (for adrq−). All sequences which encoded adw2 (genotype A) had a 2-codon insertion corresponding to amino acid 153 and 154 in the core gene, independent of origin. The most divergent residues were amino acids 74 and 91 with five and three possible variant amino acids respectively. Table 6 shows alignment of all amino acid sequences between the countries studied. Most amino acid substitutions involved the HLA-restricted CTL [3], T helper and B cell [7] epitopes, irrespective of HBV subtype (Table 5).

Genotype A (adw2 subtype) and D (ayw2) sequences had five amino acid substitutions in common at residues 12, 27, 67, 83 and 91, which were not present in other strains. Distinguishing adw2 (genotype A) and ayw2 (genotype D) was possible by examination of residues 74, 87, 97 and 116 (Table 5). At the amino acid level, the core gene sequences belonging to adw2, awy1 (both belonging to genotype B) and adrq+ (genotype C) did not show any unique amino acid substitutions (however, see nucleotide results).

Table 3 Identification of core sequences obtained from databases for comparison. Each sequence is listed under accession number

Accession no.	Identification	Origin	Subtype	Genotype	Reference
M57663	pFDW294	Philippines	adw2	A	[41]
X02763	PHBV3200	USA	adw2	A	[42]
X51970	HBV991	Germany	adw2	A	Koechel et al., unpublished data
X70185	HBV-A938	Germany	adw2	A	[43]
AF297624	Isolate 656	Africa	adw2	A	[44]
V00866	PHBV933	USA	adw2	A	[16]
Z35717	pHB614	Poland	adw2	A	Pluciennicak et al., unpublished data
X97850	Patint-4	Greece	adw2	B	[45]
X97851	Patint-6	Greece	adw2	B	[45]
AB033554	RTB299	Indonesia	adw2	B	[44]
AB033555	PAD744	Indonesia	adw2	B	[46]
AF121249	AF121249	Vietnam	adw2	B	[47]
AF282917	HBV-B_1	China	adw2	B	Hou et al., unpublished data
AF282918	HBV-B_2	China	adw2	B	Hou et al., unpublished data
AB031261	HBV Vie A-2	Vietnam	adw2	B	[48]
AB031263	HBV Vie A-5	Vietnam	adw2	B	[48]
AB031264	HBV Vie F-1	Vietnam	adw2	B	[48]
AB031262	HBV Vie A-3	Vietnam	adr	C	[48]
AB031265	HBV Vie F-2	Vietnam	adr	C	[48]
AB033553	SK619	Indonesia	adw	C	[49]
D50489	HPBA11A	Japan	adr	C	[50]
L08805	HPBETNC	Japan	adr	C	[51]
D16666	HPBE88A	Japan	adr	C	[50]
S75184	S75184	Japan	adr	C	[53]
AF241411	8290	Vietnam	adr	C	[46]
AB031260	HBV VieA-1	Vietnam	adr	C	[47]
X52939	HBV prex	China	adr	C	[53]
M38454	pADR-1	China	adr	C	[54]
M38594	M38594	Korea	adr	C	[55]
M38636	pHBV107	Korea	adr	C	[56]
X04615	pYRB259	Japan	adr	C	[18]
X14193	pADRm	Korea	adr	C	[57]
D12980	SRADR	Japan	adr	C	[58]
D00630	pHBV330	Japan	adr	C	[17]
X01587	pHBVadr4	Japan	adr	C	[59]
D00331	pAK66	Japan	adr	C	[11]
X75656	HHVCCHA	Polynesia	adrq−	C	[60]
X75665	HHVBC	New Caledonia	adrq−	C	[60]
AB033557	pIWK146	Indonesia	adw	C	[11]
AF121240	HBV/94–11066	Vietnam	ayw	D	[46]
AF121242	HBV/98–1218	Vietnam	ayw	D	[46]
AB033558	JYW796	Japan	ayw	D	[11]
AB033559	JYW310	Papua New Guinea	ayw	D	[11]
X65258	HBVAYWCI	Italy	ayw	D	Lai et al., unpublished data
X59795	HBVAYWMCG	Italy	ayw	D	[61]
X02496	pHBV320	Latvia	ayw	D	[62]
AJ131956	AJ131956	Germany	ayw	D	[63]
L27106	HPBMUT	Israel	ayw	D	[64]
X72702	HBVORFS	Germany	ayw	D	[65]
X85254	HBVPRESS12	Italy	ayw	D	Lai et al., unpublished data

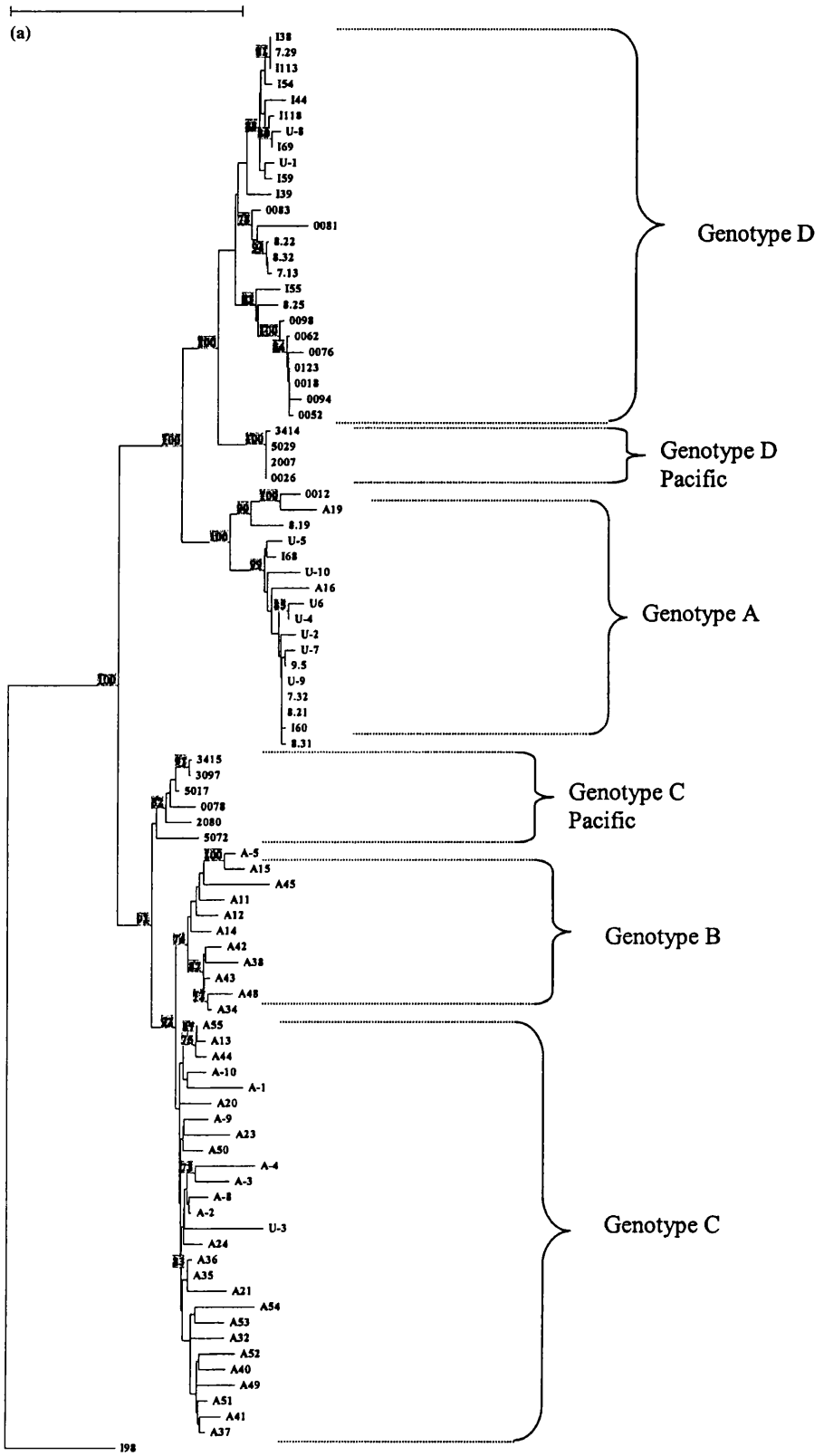


Fig. 1 Neighbour-joining phylogenetic trees of core (a) and surface (b) gene sequences from 91 HBcAg, rooted with sample I98 and U8 respectively. The figure shows bootstrap values of ≥70% and scale denotes percentage diversity. Coding numbers indicate samples that have been analysed in the figure.

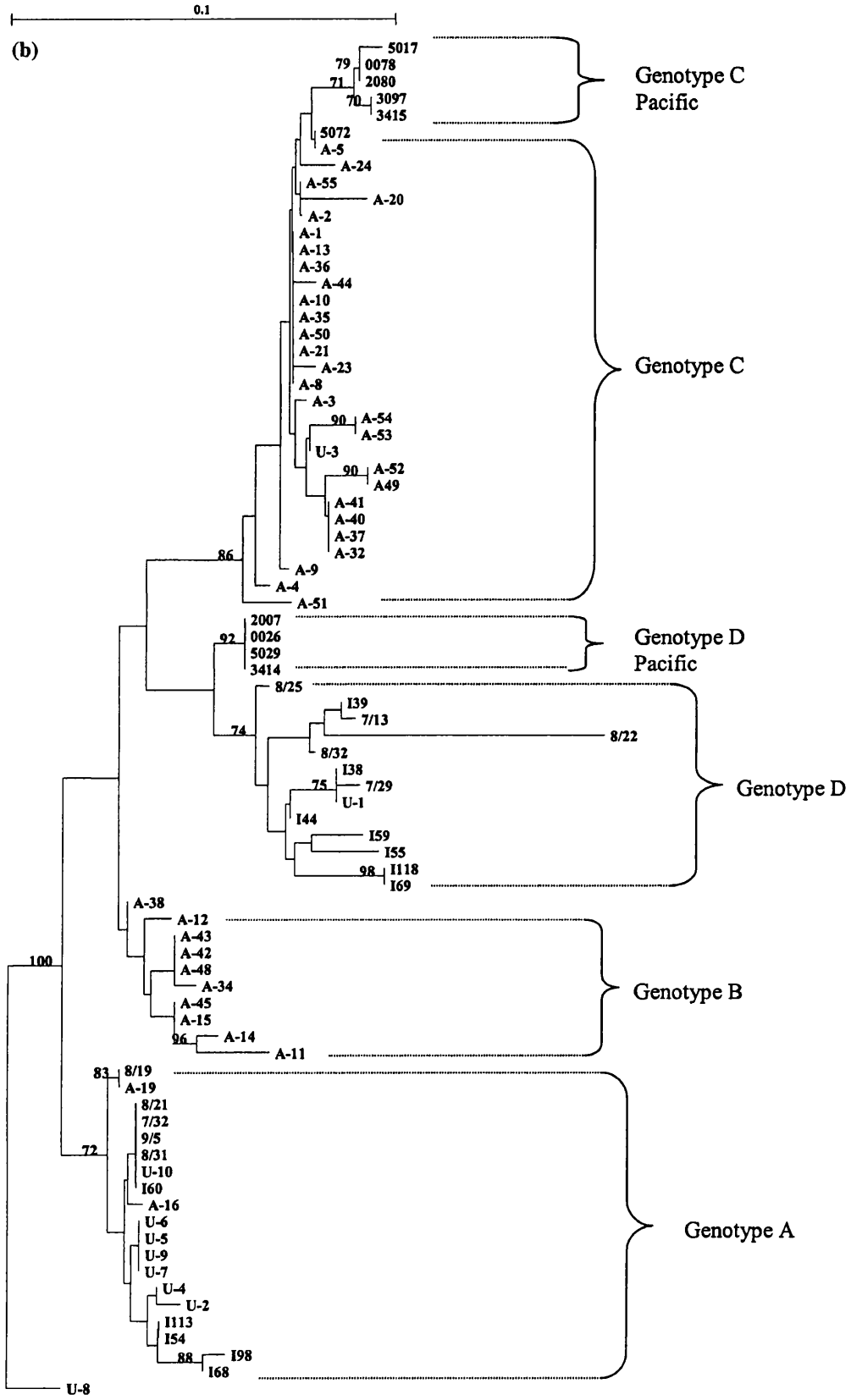


Fig. 1 Continued

Table 4 Unique C-gene nucleotide differences which permit subtype classification. Nucleotides are represented by single letters and numbered from the beginning of *Eco*RI. Italicized bold letters represent missense mutations. Underlined numbers indicate the locations of the missense nucleotides

Nucleotide position	Subtype (genotype)						
	adw2(A)	adw2(B)	ayw1(B)	adr(C)	adrq−(C)	ayw2(D)	ayw3(D)
1966	C	C	C	C	T	C	C
1975	T	T	T	T	G	T	T
1981	***C***	***T***	***T***	***T***	***T***	***A***	***A***
2059	T	G	G	A	A	T	T
2074	C	T	T	T	T	A	A
2075	***A***	***A***	***A***	***A***	***G***	***A***	***A***
2080	C	G	G	G	G	T	T
2104	A	A	A	G	G	A	A
2121	***A***	***G***	***G***	***G***	***G***	***T/G/C***	***T***
2149	***T***	***A***	***A***	***A***	***A***	***C***	***C***
2160	***A***	***G***	***G***	***G***	***G***	***G***	***G***
2189	***A***	***A***	***A***	***A/C***	***A***	***T***	***T***
2224	T	T	T	T	T	C	C
2242	T	T	T	T	T	G/A	C/A
2246	***C***	***C***	***C***	***C***	***C***	***A***	***A***
2290	A	T	T	C	C	A	A
2293	C	A	A	T	T	T	T
2296	T	T	T	C	C	T	T
2304	C	C	A	C	C	C	C
2354	C	A	C	C	C	C	C
2354–2359 insertion	CGGGAC	–	–	–	–	–	–

Table 5 The most frequent differences in HBc amino acid sequences between subtypes. Each was found in a majority of strains from each subtype and the combination was unique for that subtype. Numbering of amino acid in bold is from the beginning of HBcAg. CD4 = CD4 recognized epitope. CD8 = CD8 recognized epitope

Subtype/ genotype	CD8	CD4				CD8			
	27	59	67	74	83	87	91	97	116
adw_2/A	V	I	T	N	D	N	T	I	L
adw_2/B	I	I	N	S	E	S	V	I	L
ayw_1/B	I	I	N	S	E	S	V	I	L
adr/C	I	I	N	S	E	S	V	I	L
adrq−/C	I	V	N	S	E	S	V/I	I	L
ayw_2/D	V	I	T	G/T/V	D	S	T	F	L/I
ayw_3/D	V	I	T	V	D	S	T	F	I

Interestingly, with the exception of two samples, all strains with 97L (instead of I) were HBeAg negative or had the precore stop codon mutation, A_{1896}.

Country-specific changes

A specific motif, consisting of S12, I27, N67, S74, E83 and V91, was observed in a majority of samples from South-east Asia and the Pacific irrespective of HBV genotype, but not in core sequences from other regions studied (Table 6). Similarly, Western-derived sequences (USA-Caucasian, Scotland and Italy) shared a specific motif at the same residues, but with different amino acids: T12, V27, T67, N/V74, D83 and T91 (Table 6). This pattern seemed slightly different at amino acid position 74 in which all variation depended on genotype: N74 in genotype A and V74 in genotype D.

This finding was strengthened by analysis of the nucleotide sequence. In both populations, 19 unique nucleotide changes were found, of which eight were missense changes for six amino acid residues (results not shown).

There were nine Indian and nine Italian strains of genotype D, subtype ayw2. One sequence (represented by six examples) dominated in each country; the remaining three sequences were represented by one example each (Table 6). Further, the five Scottish strains of genotype D, subtype ayw2 or ayw3, were identical to the majority strains in India or Italy, while all the Pacific-derived strains of genotype D were unique and identical.

Seven amino acid positions allowed definition of a particular country's sequence (Table 6). In the Pacific all sequences belonging to ayw2 contained A35, D40, V59 and T74; I91 was observed in all adrq−strains from that region. I80 was seen in Indian sequences and G87 in Korea. The most heterogeneity was seen in ayw2 strains derived from Indian, Italian, Pacific island and Scottish patients. Residue 74 showed the most discrepancy between countries in which

Table 6 Alignment of complete amino acid sequences of HBcAg which shows genotype/subtype identification and geographical origins of 91 sera using Bioedit software (Clustal W). Amino acid residues are numbered from the beginning of the HBcAg using the single-letter code. Apart from genotype A, which had 185 amino acids, all the other genotypes had 183 amino acids, therefore two extra amino acids at positions 153 and 154 were added for a proper alignment

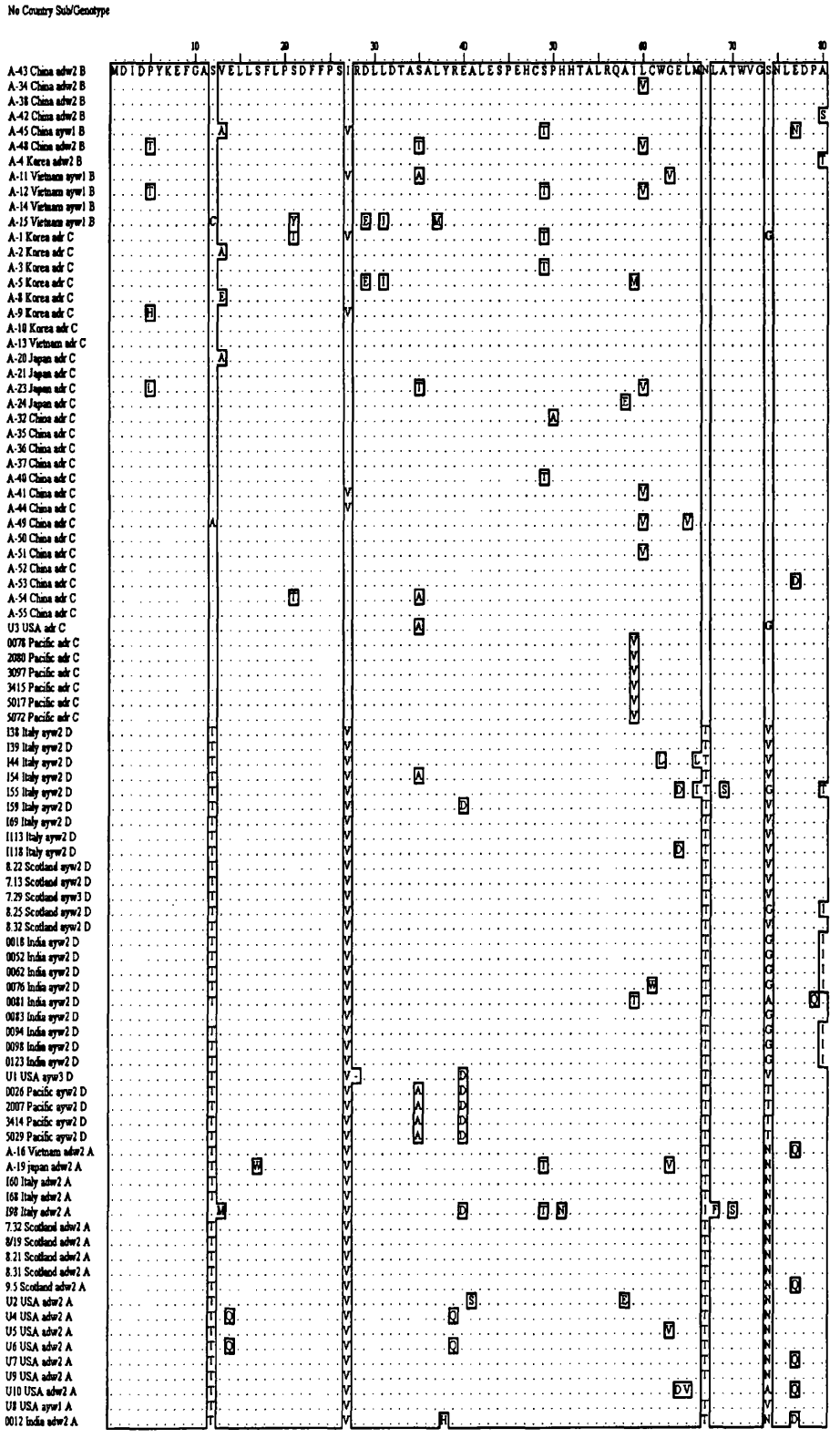


Table 6 Continued

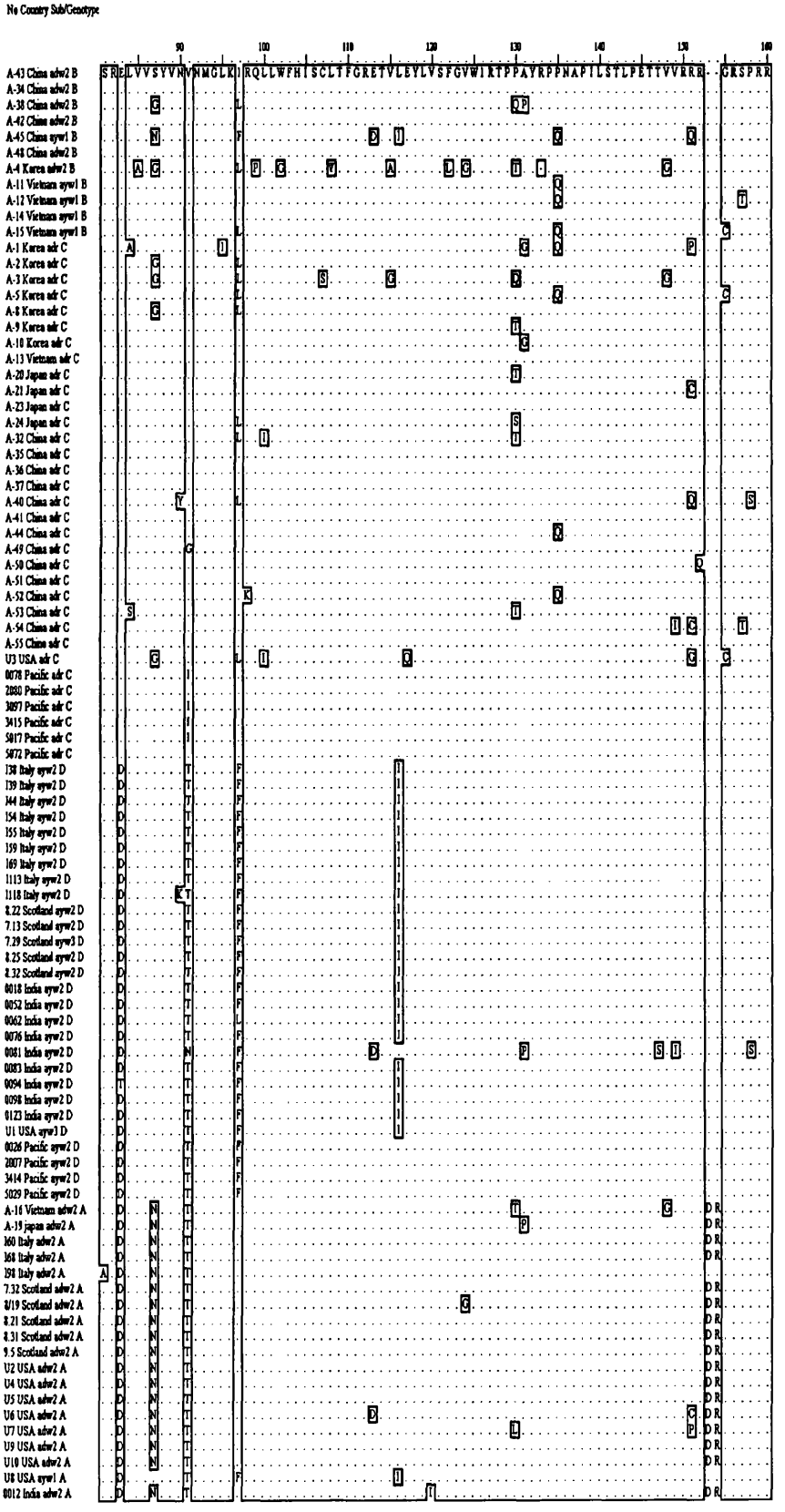
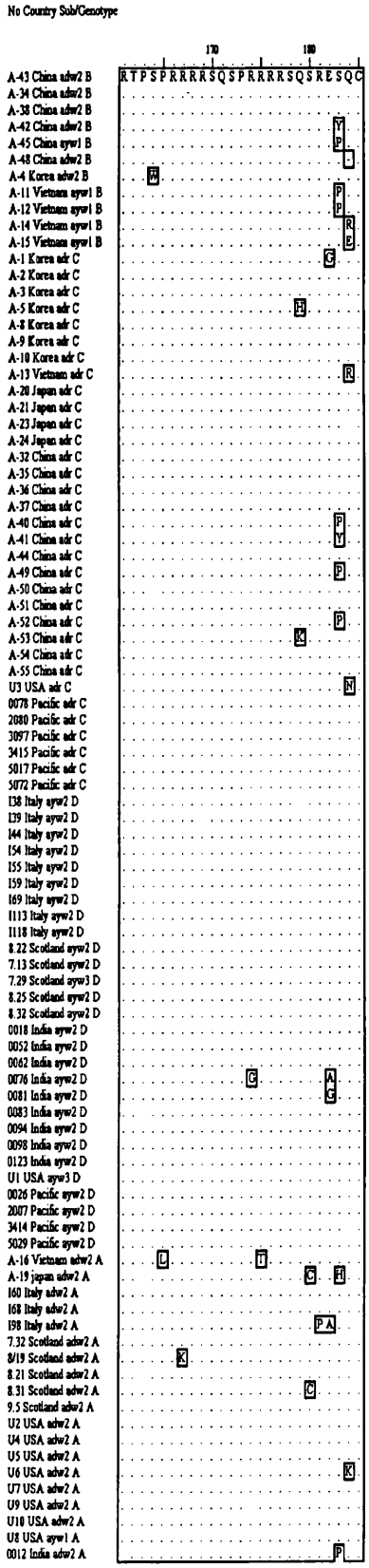


Table 6 Continued



ayw2 is common: T in Pacific, V in Scottish and Italian, and G in Indian.

However, there were additional synonymous nucleotide substitutions that were unique for specific countries in our study (results not shown). For example, in a majority of samples from India, we observed T_{1912}, C_{1955}, T_{1957}, A_{2011}, G_{2032}, A_{2138}, G_{2155} and T_{2191}. In all strains belonging to ayw2 from the Pacific region, we observed T_{1966}, G_{1975}, G_{2003}, G_{2005}, T_{2020}, A_{2089}, C_{2121} and T_{2155}. A_{2242} and C_{2189} were unique for Italian and Korean sequences respectively.

Finally, subtype-related specific nucleotide variations were observed in a majority of samples from specific countries (results not shown). As for the amino acid sequence, ayw2 subtype contained the most nucleotide divergence: 16 nucleotide substitutions were unique in ayw2 sequences from India, 10 from Pacific, and two from Italy. Much less heterogeneity was observed in adr strains: only A_{2158} and C_{2189} were unique in adr strains from Chinese and Korean sequences respectively (results not shown).

DISCUSSION

HBcAg contains helper T cell [26] and CTL epitopes [3,5,27]. Different HLA-restricted CTL/T helper epitopes have been identified within the core gene which are capable of inducing significant T-cell responses in HBV-infected patients [2–3,19]. Previous studies have suggested that epitopes for CTL/Th recognition might be different on account of the diverse distribution of HLA antigens in different geographical regions [28–30]. Since the worldwide distribution of HBV follows a geographical pattern [16], there may be a strong influence of ethnic background, perhaps driven by T-cell selection, on this distribution, reflected by divergence of amino acid substitutions within certain regions of the core gene [4,31].

The aim of this study was to characterize the core antigen variability in diverse geographical regions and ethnic groups, and to determine genotype- and subtype-specific variants. Our comparison of 91 new complete core genes and those in databases verified that amino acid/nucleotide-specific substitutions correlate with both ethnicity and HBV genotypes/subtypes. Motifs and specific individual mutations were observed that correlated with broad ethnic background, country, genotype and subtype. Our finding that HBcAg variability was distributed geographically was entirely consistent with Norder's study of S genes [16]. Moreover, construction of phylogenetic trees (Fig. 1), of C and S genes, showed an almost identical pattern of HBV genotype distribution. Some samples had discrepant genotype allocation based on analysis of two genes: in C gene samples I54 and I113 were in D, but A in S gene; similarly, sample A-5 was genotype B in C gene and C in surface gene. These discrepancies were not surprising, as recombination events between different genotypes have been described

previously [32]; between A and C from Vietnam [33], between C and D from Tibet [34] and between B and C from different countries of South-east Asia (with exception of Japan) [35]. Of interest is that A-5 came from Korea, where HBV genomes have been reported with recombination between genotypes B and C, designated as genotype Ba (a for Asia). This is in keeping with Suguachi's study [35]. Interestingly, in this study recombination between A and D was found in two of Italian sequences (I54 and I113). The validity of our data was strengthened by subsequent comparisons with international databases (Table 3). Below, we briefly discuss three aspects of our findings.

The HBV antigenic subtypes generally reflect the country or regions of origin [36]. However, there are a number of confounding factors. First, some subtypes are not geographically localized but are worldwide. Secondly, there are differences in HBV strains between native and foreign carriers in different geographical areas [36]; this means that genetically homogeneous HBV strains cannot simply be presumed to be linked epidemiologically. Thirdly, some subtypes are genetically heterogeneous and belong to more than one genotype [12]. Knowing the degree of HBV variation to be expected in a certain community could be useful in epidemiological investigations [37].

Comparisons of our data with HBcAg sequences in databases (except ayw1 strains, of which there were no examples) were almost entirely consistent. In previous reports, codon 97 showed the highest variation [38,39]; however, here, codon 74 contained the most discrepancy between sequences. Of nine variable residues observed throughout the C gene, seven were unique for genotypes A and D, and two residues were unique for the adrq– subtype. As some subtypes belong to more than one genotype, these variants may also be found in more than one genotype. Further identical variants allowed correct subtype categorization.

The findings at the amino acid level were mirrored at the nucleotide level. At the nucleotide level (Table 4), the best correlation between our data and those obtained from databases was found in genotypes A (adw2 subtype) and D strains. Because six of eight adw2 (genotype B) sequences in our study were derived from Chinese patients, we compared these with other Chinese adw sequences in the database [5]. These were almost identical. adrq+ strains in the databases were also very similar to ours, but adrq– strains did not contain T_{1975}. Tables 4 and 5 propose amino acid and nucleotide substitutions that are critical sites for defining genotypes and their corresponding subtypes.

It is interesting that peoples of a broad ethnic background (Asian or Western) share amino acid substitutions in the core gene despite having different subtypes. This was particularly true of samples from the South-east Asian regions. It would appear that there are shared host-factors, most likely immunological, which act on these sequences despite their background subtype. One of our hypotheses was that specific ethnic groups would maintain the virus that originated in their home country. The hypothesized migration of South-east Asian peoples out into the Pacific correlates with their commonality of HBV sequences; similar to Norder' study [16], six adrq– strains in genotype C as well as four ayw2 in genotype D formed a separate cluster that belonged only to the Pacific area (Fig. 1). Thus, the divergence of HBV genomes in this area is limited, in keeping with the known low variability between genotypes B and C. Another novel finding is that a majority of Indian genotype D strains almost composed a separate branch in the core phylogenetic tree (Fig. 1a). Further, we were able to distinguish strains that were observed only in Korea (Table 6). Asian immigrants to the USA shared identical sequences to those from their home country. It thus appears that there are no signature sequences in the USA, probably because of long-term, widespread importation of sequences. Immigration to the USA had no effect on Asian sequences, in keeping with circulation of HBV amongst ethnic minorities or the likelihood that, in an intra-ethnic sexual partnership, it is the Asian partner who is more likely to be the primary infecting source. Nevertheless, genotype/subtype-specific variants were identical whatever the geographical origin.

The pattern in Western-derived sequences (including from US-Caucasians, Scotland and Italy) was slightly different, as there was no specific substitution that correlated with individual countries (Table 6). Nevertheless, a clear motif was identified that was common to Western populations. We did not study genotype F, which is native to the Americas.

Core variation was most likely to occur at positions within known epitopes which can tolerate naturally occurring variants [2,4,6,20–21,38–39], i.e. residues 12, 27, 67, 74, 83, 87, 91, 97 and 116 (Table 5). Most changes were synonymous. This might be due to a lack of a positive selection advantage or a strong bias towards maintaining that amino acid. The similarity in amino acid/nucleotide distribution in adw2 (genotype A) and ayw2 (genotype D) strains found in Western countries (USA-Caucasian, Italy, and Scotland), showed that considerable constraints must exist against HBV variability in a particular genotype infecting a person of a particular background. In addition, the 'South-east Asian' and 'Western' motifs suggest either that there is positive selection on specific variants or that genetic drift in HBV is relatively slow [40]. According to the almost unique amino acid variation in residues 35, 40, 59, 74, 80, 87 and 91 in particular countries (Table 6), we hypothesize that differences in distribution of HLA antigens, or other immune genes between diverse geographical areas [6], probably contributed to the selection of amino acid variation. However, primary data need to be accumulated to investigate this issue.

In conclusion, it appears from this work that there are three levels at which C gene variation relates to ethnic background. First is that there are various genotypes associated with specific subtypes, which may have migrated with their human hosts. Next, peoples with related ethni-

city, despite wide geographical location, have a set of common variants. Finally, small changes (one or two substitutions) are selected, either negatively or positively, once there is little intermixing between these people (e.g. once they became a tribe). As more sequence data accumulate from different geographical origins, it is expected that HBV genomic classification will become refined, contributing to the finer mapping of the relation between diverse geographical origins and the distribution of HBV strains. Further, sequence analysis of core gene, particularly, in the CTL-restricted regions, provides a useful tool for better understanding the biology of HBV strains, and finally, for design of immunomodulatory therapies upon vaccination.

ACKNOWLEDGEMENTS

We wish to thank Dr Helen Norder for her helpful advice and reading and criticism of the manuscript. We are grateful to Dr David Yirrell for carrying out the phylogenetic analysis.

REFERENCES

1 Hur GM, Lee YI, Suh DJ, Lee JH, Lee YI. Gradual accumulation of mutations in precore core region of HBV in patients with chronic active hepatitis: implication of clustering changes in a small region of the HBV core region. *J Med Virol* 1996; 48: 38–46.

2 Ferrari C, Berlotti A, Penna A *et al.* Identification of immunodominant T cell epitopes of the hepatitis B virus necleocapsid antigen. *J Clin Invest* 1991; 88: 214–222.

3 Ehata T, Omata M, Yokosuka O, Hosoda K, Ohto M. Variations in codons 84–101 in the core nucleotide sequence correlate with hepatocellular injury in chronic hepatitis B virus infection. *J Clin Invest* 1992; 89: 332–338.

4 Bertoletti A, Ferrari C, Fiaccadori F *et al.* HLA class I-restricted human cytotoxic T cells recognize endogenously synthesized hepatitis B virus nucleocapsid antigen. *Proc Natl Acad Sci USA* 1991; 88: 10445–10449.

5 Chuang WL, Omata M, Ehata T *et al.* Precore mutations and core clustering mutations in chronic hepatitis B virus infection. *Gastroenterology* 1993; 104: 263–271.

6 Ehata T, Omata M, Chuang WL *et al.* Mutations in the core nucleotide sequence of hepatitis B virus correlated with severe liver damage. *Viral Hepat Liver Dis* 1994; ??: 277–282.

7 Carman WF, Thursz M, Hadziyannis S *et al.* Hepatitis B e antigen negative chronic active hepatitis: hepatitis B virus core mutations occur predominantly in known antigenic determinants. *J Viral Hepat* 1995; 2: 77–84.

8 Okumura A, Takayanagi M, Aiyama T *et al.* Serial analysis of hepatitis B virus core nucleotide sequence of patients with acute exacerbation during chronic infection. *J Med Virol* 1996; 49: 103–109.

9 Caillat-Zucman S, Gimenez JJ, Wambergue F *et al.* Distinct HLA class II alleles determine antibody response to vaccination with hepatitis B surface antigen. *Kidney Int* 1998; 53(6): 1626–1630.

10 Ahn SH, Han KH, Park JY *et al.* Association between hepatitis B virus infection and HLA-DR type in Korea. *Hepatology* 2000; 31(6): 1371–1373.

11 Okamoto H, Tsuda F, Sakugawa H *et al.* Typing hepatitis B virus by homology in nucleotide sequences: comparison of surface antigen subtypes. *J Gen Virol* 1988; 69: 2575–2583.

12 Norder H, Courouce AM, Magnius LO. Molecular basis of hepatitis B virus serotype variations within the four major subtypes. *J Gen Virol* 1992; 73: 3141–3145.

13 Stuyver L, De Gendt S, Van Geyt C *et al.* A new genotype of hepatitis B virus: complete genome and phylogenetic relatedness. *J Gen Virol* 2000; 81: 67–74.

14 Arauz-Ruiz P, Norder H, Robertson BH, Magnius LO. Genotype H: a new Amerindian genotype of hepatitis B virus revealed in Central America. *J Gen Virol* 2002; 83(Pt 8): 2059–2073.

15 Courouce-Pauty AM, Lemaire JM, Roux JF. New hepatitis B surface antigens subtypes inside the ad category. *Vox Sang* 1978; 35: 304–308.

16 Norder H, Hammas B, Lee SC *et al.* Genetic relatedness of hepatitis B viral strains of diverse geographical origin and natural variations in the primary structure of the surface antigen. *J Gen Virol* 1993; 74: 1341–1348.

17 Ono Y, Onda H, Sasada R, Igarashi K, Sugino Y, Nishioka K. The complete nucleotide sequences of the cloned hepatitis B virus DNA; subtype adr and adw. *Nucleic Acids Res* 1983; 25: 1747–1757.

18 Okamoto H, Imai M, Shimozaki M *et al.* Nucleotide sequence of a cloned hepatitis B virus genome, subtype ayr: comparison with genomes of the other three subtypes. *J Gen Virol* 1986; 67: 2305–2314.

19 Ehata T, Omata M, Chuang WL *et al.* Mutations in core nucleotide sequence of hepatitis B virus correlate with fulminant and severe hepatitis. *J Clin Invest* 1993; 91: 1206–1213.

20 Akarca US, Lok AS. Naturally occurring hepatitis B virus core gene mutations. *Hepatology* 1995; 22: 50–60.

21 Hosono S, Tai PC, Wang W *et al.* Core antigen mutations of human hepatitis B virus in hepatomas accumulate in MHC class II-restricted T cell epitopes. *Virology* 1995; 10: 151–162.

22 Hall TA. BioEdit: a user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT. *Nucleic Acids Symp Ser* 1999; 41: 95–98.

23 Van de Peer Y, De Wachter, R. TREECON for Windows: a software package for the construction and drawing of evolutionary trees for the Microsoft Windows environment. *Comput Appl Biosci* 1994; 10: 569–570.

24 Kimura M. A simple method for estimating evolutionary rates of base substitutions through comparative studies of nucleotide sequences. *J Mol Evol* 1980; 16: 111–120.

25 Felenstein J. Confidence limits on phylogenies: an approach using the bootstrap. *Evolution* 1985; 39: 783–791.

26 Jung MC, Diepolder HM, Sepngler U *et al.* Activation of a heterogeneous hepatitis B (HB) core and e antigen-specific CD4+ T-cell population during seroconversion to anti-HBe and anti-HBs in hepatitis B virus infection. *J Virol* 1995; 69: 3358–3368.

27 Wakita T, Kakuma S, Shibata M *et al.* Detection of pre-C and core region mutants of hepatitis B virus in chronic hepatitis B virus in chronic hepatitis B virus carriers. *J Clin Invest* 1991; 88: 1793–1801. Properties of hepatitis B virus genome recovered from Vietnamese patients with fulminant hepatitis in comparison with those of acute hepatitis. *J Med Virol* 2000; 61: 23–28.

28 Thursz MR, Kwiatkowski D, Allsopp CE, Greenwood BM, Thomas HC, Hill AV. Association between an MHC class II allele and clearance of hepatitis B virus in the Gambia. *N Engl J Med* 1995; 332(16): 1065–1069.

29 Hohler T, Gerken G, Notghi A *et al.* HLA-DRB1*1301 and *1302 protect against chronic hepatitis B. *J Hepatol* 1997; 26(3): 503–507.

30 Thio CL, Carrington M, Marti D *et al.* Class II HLA alleles and hepatitis B virus persistence in African Americans. *J Infect Dis* 1999; 179(4): 1004–1006.

31 Zampino R, Lobello S, Chiaramonte M *et al.* Intra-familial transmission of hepatitis B virus in Italy: phylogenetic sequence analysis and amino-acid variation of the core gene. *J Hepatol* 2002; 36(2): 248–253.

32 Morozov V, Pisareva M, Groudinin M. Recombination between different genotypes of hepatitis B virus. *Gene* 2000; 260(1–2): 55–65.

33 Hannoun C, Horal P, Lindh M. Long-term mutation rates in the hepatitis B virus genome. *J Gen Virol* 2000; 81(Pt 1): 75–83.

34 Cui C, Shi J, Hui L *et al.* The dominant hepatitis B virus genotype identified in Tibet is a C/D hybrid. *J Gen Virol* 2002; 83(Pt 11): 2773–2777.

35 Sugauchi F, Orito E, Ichida T *et al.* Hepatitis B virus of genotype B with or without recombination with genotype C over the precore region plus the core gene. *J Virol* 2002; 76(12): 5985–5992.

36 Courouce-Pauty AM, Plancon A, Soulier JP. Distribution of HBsAg subtypes in the world. *Vox Sang* 1983; 44: 197–211.

37 Bläckberg J, Kidd-Ljunggren ??. Occult hepatitis B virus after acute self-limited infection persisting for 30 years without sequence variation. *J Hepatol* 2000; 33: 992–997.

38 Yuan TT, Tai PC, Shih C. Subtype-independent immature secretion and subtype-dependent replication deficiency of a highly frequent, naturally occurring mutation of human hepatitis B virus core antigen. *J Virol* 1999; 73: 10122–10128.

39 Yuan TT, Shih C. A frequent, naturally occurring mutation (P130T) of human hepatitis B virus core antigen is compensatory for immature secretion phenotype of another frequent variant (I97L). *J Virol* 2000; 74: 4929–4932.

40 Wilson JN, Nokes DJ, Carman WF. The predicted pattern of emergence of vaccine-resistant hepatitis B: a cause for concern? *Vaccine* 1999; 17(7–8): 973–978.

41 Estacio RC, Chavez CC, Okamoto H *et al.* Nucleotide sequence of a hepatitis B virus subtype adw isolated from a Filipino: comparison with the three reported genomes of the same subtype. *J Gastroenterol Hepatol* 1988; 3: 214–222.

42 Valenzuela P, Quiroga M, Zaldivar J, Gray P, Rutter WJ. The nucleotide sequence of the hepatitis B viral genome and the identification of the major viral genes. In: Fields B, Jaenisch R, Fox C, eds. Animal Virus Genetics. New York: Academic Press, 1980: 57–70.

43 Preisler-Adams S, Schlayer HJ, Peters T, Korp R, Rasenack J. Complete nucleotide sequence of a hepatitis B virus, subtype adw2, and identification of three types of C open reading frame. *Nucleic Acids Res* 1993; 11: 2258.

44 Owiredu WK, Kramvis A, Kew MC. Molecular analysis of hepatitis B virus genomes isolated from black African patients with fulminant hepatitis B. *J Med Virol* 2001; 65(3): 485–492.

45 Alexopoulou A, Karayiannis P, Hadziyannis SJ *et al.* Whole genome analysis of hepatitis B virus from four cases of fulminant hepatitis: genetic variability and its potential role in disease pathogenicity. *J Viral Hepat* 1996; 3(4): 173–181.

46 Sastrosoewignjo RI, Okamoto H, Mayumi M, Warsa UC, Sujudi X. The complete nucleotide sequence of an HBV DNA clone of subtype adw (pRTB299) from Indonesia. *ICMR Ann* 1985; 5: 39–40.

47 Sastrosoewignjo RI, Okamoto H, Mayumi M, Rustam M, Warsa UC, Sujudi X. The complete nucleotide sequence of an HBV DNA clone of subtype adw (pPAD744) from Sumatra, Indonesia. *ICMR Ann* 1986; 6: 99–106.

48 Hannoun C, Norder H, Lindh M. An aberrant genotype revealed in recombinant hepatitis B virus strains from Vietnam. *J Gen Virol* 2000; 81: 2267–2272.

49 Okamoto H, Imai M, Kametani M, Nakamura T, Mayumi M. Genomic heterogeneity of hepatitis B virus in a 54-year-old woman who contracted the infection through materno-fetal transmission. *Jpn J Exp Med* 1987; 57: 231–236.

50 Uchida T, Gotoh K, Shikata T. Complete nucleotide sequences and the characteristics of two hepatitis B virus mutants causing serologically negative acute or chronic hepatitis B. *J Med Virol* 1995; 45: 247–252.

51 Ogata N, Miller RH, Ishak KG, Purcell RH. The complete nucleotide sequence of a pre-core mutant of hepatitis B virus implicated in fulminant hepatitis and its biological characterization in chimpanzees. *Virology* 1993; 194: 263–276.

52 Moriyama K, Takada T, Tsutsumi Y *et al.* Mutations in the transcriptional regulatory region of the precore and core/pregenome of a hepatitis B virus with defective HBeAg production. *Fukuoka Igaku Zasshi* 1994; 85: 314–322.

53 Locarevic IF, Zentgraf H, Schroder CH. Sequence of a replication competent hepatitis B virus genome with a pre-X open reading frame. *Nucleic Acid Res* 1990; 18: 4940.

54 Renbao G, Meijin C, Lueping S, Suwen Q, Zaiping L. The complete nucleotide sequence of the cloned DNA of hepatitis B virus subtype adr in pADR-1. *Sci Sin B* 1987; 30: 507–521.

55 Kim YS, Hyun SW, Rho HM. Nucleotide sequence of the core-antigen coding region of hepatitis B virus (subtype adr-k). *Korean J Biochem* 1985; 18: 304–312.

56 Kim KT, Hyun SW, Kim YS, Rho HM. Nucleotide sequence of the core-antigen coding region of hepatitis B virus (subtype adr-k). *Korean J Biochem* 1988; 18: 319–331.

57 Rho HM, Kim K, Hyun SW, Kim YS. The nucleotide sequence and reading frames of a mutant hepatitis B virus subtype adr. *Nucleic Acids Res* 1989; 11: 2124.

MAJOR ARTICLE

Intracellular Distribution of Hepatitis B Virus Core Protein Expressed In Vitro Depends on the Sequence of the Isolate and the Serologic Pattern

Mohammed S. Jazayeri,[1] Edward S. Dornan,[1] Winifred Boner,[1] Giovanna Fattovich,[3] Stephanos Hadziyannis,[4] and William F. Carman[1,2]

[1]Division of Virology, IBLS, University of Glasgow, and [2]West of Scotland Specialist Virology Centre, Gartnavel General Hospital, Glasgow, Scotland; [3]Servizio Autonomo Clinicizzato di Gastroenterology, University of Verona, Verona, Italy; [4]Department of Medicine, Hippokration Hospital, Athens, Greece

Intracellular localization of hepatitis B core antigen (HBcAg) in vivo varies with liver cell damage. By use of confocal microscopy, in vitro localization of HBcAg was studied using transfection of cloned HBcAg variants in Cos-7 cells, by use of the pRK5 expression vector. Twenty-six samples were obtained from 14 patients with liver disease; 10 of the samples were hepatitis B e antigen (HBeAg) positive, and 16 were anti–hepatitis B e (HBe) positive. In HBeAg–positive patients, HBcAg predominantly localized in the nucleus; in anti-HBe–positive patients, it accumulated mainly in the cytoplasm. This finding also applied to patients who either experienced seroconversion to anti-HBe or remained anti-HBe positive and developed mutations. Of the 13 samples with nuclear localization, 9 were HBeAg positive; 5 had C-terminus and/or B cell epitope mutations. All but 1 of the 13 samples with predominantly cytoplasmic localization were anti-HBe positive; all 13 had mutations. Using site-directed mutagenesis to revert C-terminus and B cell epitope mutant sequences with cytoplasmic expression back to the wild type led to the shifting of HBcAg distribution back to a predominantly nuclear distribution. Thus, the pattern of HBcAg localization in vitro depends on sequence and the serologic pattern of chronic infection, paralleling the situation in vivo.

The hepatitis B nucleocapsid protein consists of 2 structurally and functionally separate domains; the sequence from the N-terminus to approximately amino-acid (aa) position 149 is, by itself, sufficient for assembly, whereas the C-terminus is involved in pregenomic encapsidation. Hepatitis B core antigen (HBcAg) is a major immune target [1, 2] that can function as a T cell–dependent and T cell–independent antigen [3]. Hepatitis B e antigen (HBeAg), on the other hand, is a nonessential secreted variant of the core protein, and it is believed to be an immune modulator that somehow suppresses cytotoxic T lymphocyte (CTL) responses. Once an immune response against HBeAg has developed, it is advantageous for the virus to suppress production of HBeAg. Pathogenicity results from the inability to avoid reinfection of hepatocytes and from the incomplete suppression of replication by T cells. Although HBcAg and HBeAg are serologically distinct, the primary amino-acid sequence is identical over a large stretch. Three shared binding sites for anti-HBe and anti-HBc have been mapped to core aa positions 74–89, 107–118, and 128–135 of HBcAg (table 1). An additional antigenic site has been proposed within the arginine-rich C-terminus domain of HBcAg that, unlike the other epitopes, faces the interior of the nucleocapsid [4] (table 1). These 2 proteins are also highly cross-reactive at the T helper (Th) and CTL levels. In chronic carriers, seroconversion from HBeAg to anti–hepatitis B e (HBe) can lead to either ongoing, often severe, disease or remission of hepatitis. Among patients who experienced seroconversion from HBeAg to anti-HBe, there exists supportive indirect evidence for the proposed immunomodulatory role of HBeAg [5–7]. In patients whose disease went into remission, mu-

Received 23 May 2003; accepted 11 August 2003; electronically published XX April 2004.

Reprints or correspondence: Dr. William F. Carman, West of Scotland Specialist Virology Center, Gartnavel General Hospital, PO Box 16766, Glasgow, G12 0ZA, Scotland (w.carman@vir.gla.ac.uk).

The Journal of Infectious Diseases 2004;189:XXX–XXX

0022-1899/2004/18907-00XX$15.00

Table 1. Antigenic sites proposed within hepatitis B core antigen (HBcAg).

Reference, year	Epitope/HBcAg amino acid	Immune effector	No. of mutations[a]
Penna, 1992	1–20	Th	8
Bertoletti, 1991	18–27	CTL	5
Ferrari, 1991	50–69	Th	32
Salfeld, 1989	74–89	B	21
Ehata, 1992	84–101	CTL	4
Colucci, 1988	107–118	B	3
Ferarri, 1991	117–131	Th	1
Salfeld, 1989	128–135	B	9
Missale, 1993	141–151	CTL	2
Machida, 1989	148–160	B	13

NOTE. Nos. denote amino-acid residues within the core protein. B, B cell epitope; CTL, cytotoxic T lymphocyte epitope; Th, T helper epitope.

[a] For all patients in the present study.

tations tended to cluster in CD4-Th epitopes, whereas, in patients with ongoing disease, mutations clustered in B cell epitopes [8–9], although not in CTL epitopes [10–12]. Nearly all those selections occurred at the time of, or soon after, emergence of the precore stop codon. One interpretation of this finding is that the loss of HBeAg led to increased immune activity. Samples that were obtained later during the anti-HBe–positive phase, during peaks in hepatic activity, showed few additional changes, and such changes did not occur within epitopes. These and other results have led to a conundrum concerning the pathogenesis of disease in this group. If hepatitis B virus (HBV) disease is immune mediated, as is believed, then why should viruses, which appear to have escaped the immune response, still cause disease? One suggestion is that there may be immune targets in other regions of the genome, in which case one might predict that selection would occur sequentially in these other proteins. Alternatively, the mutated viruses may be directly cytopathic or may become novel targets for other relatively unstudied arms of the immune response, such as NK cells [13–15].

Several histologic studies of the association between expression of HBcAg and liver cell damage have been performed. However, little is known about cellular localization of HBcAg variants. For HBeAg-positive patients [16–20] and for anti-HBe–positive patients whose disease is in remission [18, 21, 22], staining for HBcAg in liver biopsy specimens tends to be nuclear; however, for anti-HBe–positive patients with ongoing disease [17, 23–26], staining tends to be both nuclear and cytoplasmic. This cannot be related only to a high level of replication, because HBeAg-positive patients usually are highly viremic. In vitro, the cellular localization of HBcAg is dependent on the stage of the cell cycle [27]. HBcAg is usually confined to the nucleus when cells are blocked in the G1/S stage. This may explain, in part, the in vivo observations that core protein is expressed more predominantly in the cytoplasm during periods of active disease when liver cells are dividing. Conversely, during periods of clinical remission of disease, when hepatocytes are quiescent, core protein is mainly confined to the nucleus. If sequence could be shown to correlate with distribution of HBcAg—especially if distribution changes with selection of mutations in individual patients—this would provide important clues regarding the potential mechanisms of liver damage. q14

In the present study, we examined the distribution of HBcAg in transfected cells from samples successively obtained from patients with chronic hepatitis B disease. The distribution of HBcAg was correlated with sequence variation within B cell epitopes and the C-terminus. Next, using mutagenesis, we directly investigated the influence of these mutations with regard to intracellular trafficking. q15

PATIENTS, MATERIALS, AND METHODS

Patients. All 14 patients who participated in the study were Greek or Italian, had known HBV genotypes, and were under the care of the authors of the present study. The patients were selected for inclusion in the study, on the basis of previous characterization of their sequential HBV precore/core sequences [28]. Twenty-six samples were studied and were categorized as belonging to 1 of 3 groups (table 2). Group 1 comprised all samples obtained from 6 patients. Group 1-A included only the samples that were obtained from these 6 patients when they were HBeAg positive (including 2 initial samples [$1A^1$ and $1A^2$] obtained from the first patient). Group 1-B included a second set of samples from the same 6 patients, all of which were obtained after the patients underwent anti-HBe seroconversion. Group 2 included all samples obtained from 5 patients who had active disease and who were found to be continuously anti-HBe positive. Group 2-A included the samples obtained from these 5 patients during an early period when few mutations were apparent in HBcAg. Group 2-B included a second set of samples obtained from the same 5 patients after selection of either a precore stop codon and/or at $\geq$1 HBcAg mutation. Group 3 included all samples obtained, at a single point in time, from HBeAg-positive patients. q16 q17 q18 q19

Polymerase chain reaction (PCR), cloning, and mutagenesis. DNA was extracted from 50 μL of serum, by use a QIAamp blood kit (Qiagen). Nested PCR was performed on 5 μL of extracted DNA, by use of primers to the core region of HBV, as described elsewhere [8]. One microliter of first-round PCR products was nested using primers that contained the restriction enzyme sites *Eco*RI and *Hin*dIII, respectively. The second-round PCR products were purified using Geneclean (BIO101), according to the manufacturer's instructions. Ten microliters of the products purified using Geneclean were used q20

Table 2. Serologic, clinical, and sequence data correlated with hepatitis B core antigen (HBcAg) localization in vitro.

Group,[a] sample	Genotype	HBeAg/ anti-HBe	Clinical outcome	HBcAg distribution	Mutation		
					B cell epitope[b]	C-terminus	Precore stop codon[c]
Group 1							
$1A^1$	A	+/−		N	−	−	−
$1A^2$		+/−		C	+	+	−
1B		−/+	Remission	C	+	+	−
2A	D	+/−		N	−	−	−
2B		−/+	Remission	C	−	+	+
3A	D	+/−		N	−	−	+
3B		−/+	Remission	C	+	+	+
4A	D	+/−		N	−	−	−
4B		−/+	Active	C	+	+	+
5A	D	+/−		N	+	−	−
5B		−/+	Active	C	+	+	+
6A	D	+/−		N	+	+	+
6B		−/+	Active	C	+	+	+
Group 2							
7A	D	−/+		NC	+	+	−
7B		−/+	Remission	NC	+	+	+
8A	D	−/+		N	−	+	+
8B		−/+	Active	C	+	+	+
9A	D	−/+		N	+	−	−
9B		−/+	Active	NC	+	+	+
10A	D	−/+		C	+	−	+
10B		−/+	Active	N	−	−	−
11A	D	−/+		N	+	+	+
11B		−/+	Active	C	+	+	+
Group 3							
12A	D	+/−	Acute disease	N	−	−	−
13A	A	+/−	Acute disease	N	−	−	−
14A	D	+/−	Acute disease	N	−	−	−

NOTE. C, Predominantly cytoplasmic distribution; HBe, hepatitis B e; HBeAg, hepatitis B e antigen; N, predominantly nuclear distribution; NC, mixed nuclear and cytoplasmic distribution; +, positive; −, negative.

[a] Group 1 included samples obtained from patients who had seroconversion from HBeAg to anti-HBe, group 2 included samples obtained from patients who were continuously anti-HBe positive, and group 3 included single samples obtained from patients who were HBeAg positive. Two initial samples were obtained from patient 1 (samples $1A^1$ and $1A^2$).

[b] B cell epitopes were those shown in table 1.

[c] Mutation A_{1896}.

in a restriction digest with 10 U each of *Eco*RI and *Hin*dIII (Life Technologies), were run on a 1% agarose gel, and were purified using Geneclean. These fragments were ligated into the multiple-cloning site of the eukaryotic expression vector pRK5 (a gift from H. Will, Hamburg, Germany), which contained a cytomegalovirus promoter. Competent *Escherichia coli* strain DH5α was transformed with the ligation reactions, colonies were picked, and clones that contained inserts were identified by restriction digests of minipreparations of DNA. Large-scale preparations of DNA for transfection were prepared over 2 CsCl density gradients. The constructs were sequenced to confirm the presence of mutations.

Site-directed mutagenesis was performed with a PCR method that used oligonucleotides to revert mutated sequences back to the wild-type sequence (according to original sequences in the early phase and/or normal variations in the corresponding genotype, as obtained from a database), according to the manufacturer's instructions (Stratagene). Sequencing of the plasmid confirmed the presence of mutations. All relevant sequence information is found in table 3. q23

Table 3. In vitro mutation of cytoplasmically distributed sequences.

Mutated virus	Mutagenesis position[a]			HBcAg localization		
	B cell epitope					
	HBc-1[b]	HBc-2[c]	C-terminus	N	NC	C
2B			T160A	+++	+	
3B	N74A			+++	+	
3B			S181P	++	++	
3B	N74A		S181P	+	+++	
4B	E77D, A79Q, S80P			+	+++	
4B			S181P	+++	+	
4B	E77D, A79Q, S80P		S181P	+++	+	
5B			Q177K	+++	+	
5B	A81T		Q177K	+++	+	
10A	V74A			++	++	
10A		A131P		++		++
10A	N74A	A131P		+	+++	
11B		P135Q		+	+++	

NOTE. Confocal microscopy was used to study the distribution of hepatitis B core antigen (HBcAg) after mutation was confirmed by sequencing. C, Predominantly cytoplasmic distribution; N, predominantly nuclear distribution; NC, mixed nuclear and cytoplasmic localization; +, 25% of stained cells; ++, 50% of stained cells; +++, 75% of stained cells.

[a] The mutations generated are listed according to their amino-acid position within an epitope. Letters appearing before each no. denote the mutated sequence to wild type, and letters appearing after each no. denote the original sequence before mutagenesis.

[b] Residues between aa positions 74–89.

[c] Residues between aa positions 128–135.

Cell culture, transfection, and immunofluorescence. It is known that the distribution of expressed HBcAg is cell cycle dependent. Sequential thymidine- and aphidicolin-blocking steps produced monolayers of synchronized Cos-7 cells (data not shown). Twelve hours before treatment with cell cycle–blocking agents, Cos-7 cells were seeded at a density of 3×10^4 cells/well, in a 24-well plate, in growth medium. Medium that contained 2 mmol/L thymidine replaced the aforementioned medium, and, 12 h later, cells were washed twice with PBS and refed with normal medium. Transfections were done after thymidine withdrawal and cell release, by use of FuGENE-6 and/or Lipofectamine Plus transfection reagent (Invitrogen), according to the manufacturer's instructions. After an additional 13 h, the cells were washed twice with PBS, medium that contained 2.5 μg/mL aphidicolin was added, and the cells then were incubated for an additional 36 h. Aphidicolin blocks cells in the G1/S phase of the cell cycle, in which expression of HBcAg is predominantly nuclear [27]. This inhibits movement of HBcAg into the cytoplasm because of dissolution of the nuclear membrane, implying that any cytoplasmic expression is more likely to be the result of active transport.

FACS analysis. The cell-cycle status of cultured cells was determined by flow cytometric analysis of permeabilized cells that were stained with propidium iodide. Cos-7 cells were dissociated with EDTA at the appropriate time after transfection and were resuspended in PBS. Cells underwent centrifugation at 1500 *g* for 2 min and then were fixed with 70% ethanol on ice. After 20 min, the cells underwent centrifugation, were washed once with PBS, and then underwent centrifugation again. Pelleted cells were resuspended in 1 mL of a solution that contained 0.5% Triton X-100, propidium iodide, and RNase. After incubation for 30 min at room temperature, the total DNA content was analyzed in the FL3-H channel of a FACS flow cytometer (Beckton Dickinson). q26

Immunofluorescence and confocal microscopy. Localization of core protein was studied using immunofluorescence and confocal microscopy. Cells were fixed with ice-cold methanol. After permeabilization with 0.5% Triton X-100, rabbit anti–core polyclonal Ig G (Zymed) and fluorescein isothiocyanate–conjugated goat anti–rabbit IgG (Sigma) were used as primary and secondary antibodies, respectively. Then, cell samples were examined using a Zeiss LSM510 confocal microscope with laser excitation at 488 nm. The intracellular distribution and quantitative expression of HBcAg were determined by random evaluation of ⩾100 cells with either a 40× or a 63× oil-immersion objective lens. Cells had either predominantly nuclear distribution or predominantly cytoplasmic distribution if >70%–80% of cells showed nuclear or cytoplasmic HBcAg, respectively. If a similar proportion of cells expressed HBcAg in both the nucleus and the cytoplasm, they were considered to have nuclear and cytoplasmic localization. Data sets were processed using LSM510 software. q27 q28

RESULTS

Synchronization of Cos-7 Cells by Thymidine-Aphidicolin

To study the influence of the cell cycle on the subcellular distribution of HBcAg, Cos-7 cells were synchronized by the use of 2 blocking agents. In initial experiments (data not shown), aphidicolin, when used alone soon after transfection, yielded low transfection efficiencies. Consequently, a 2-step cell-cycle synchronization was used. As a first step, cells were incubated in the presence of thymidine for 12 h, to enter the G0/G1 phase of the cell cycle (figure 1*A*). Afterward, cells were withdrawn from thymidine and were permitted to progress out of the G0/G1 phase and into the S, G2, and then M phases of the cell cycle for 12 h (figure 1*B*). After cell-cycle release, the cells were washed twice with PBS, and transfection was done using FuGENE-6 and/or Lipofectamin Plus transfection reagent methods. As a second step, 12 h after transfection, the cells were treated with 2.5 μg/mL aphidicolin for 30 h. Synchronization achieved by use of aphidicolin resulted in 94% of the cells remaining in the G1/S phase (figure 1*D* and 1*E*). q29 q30 q31

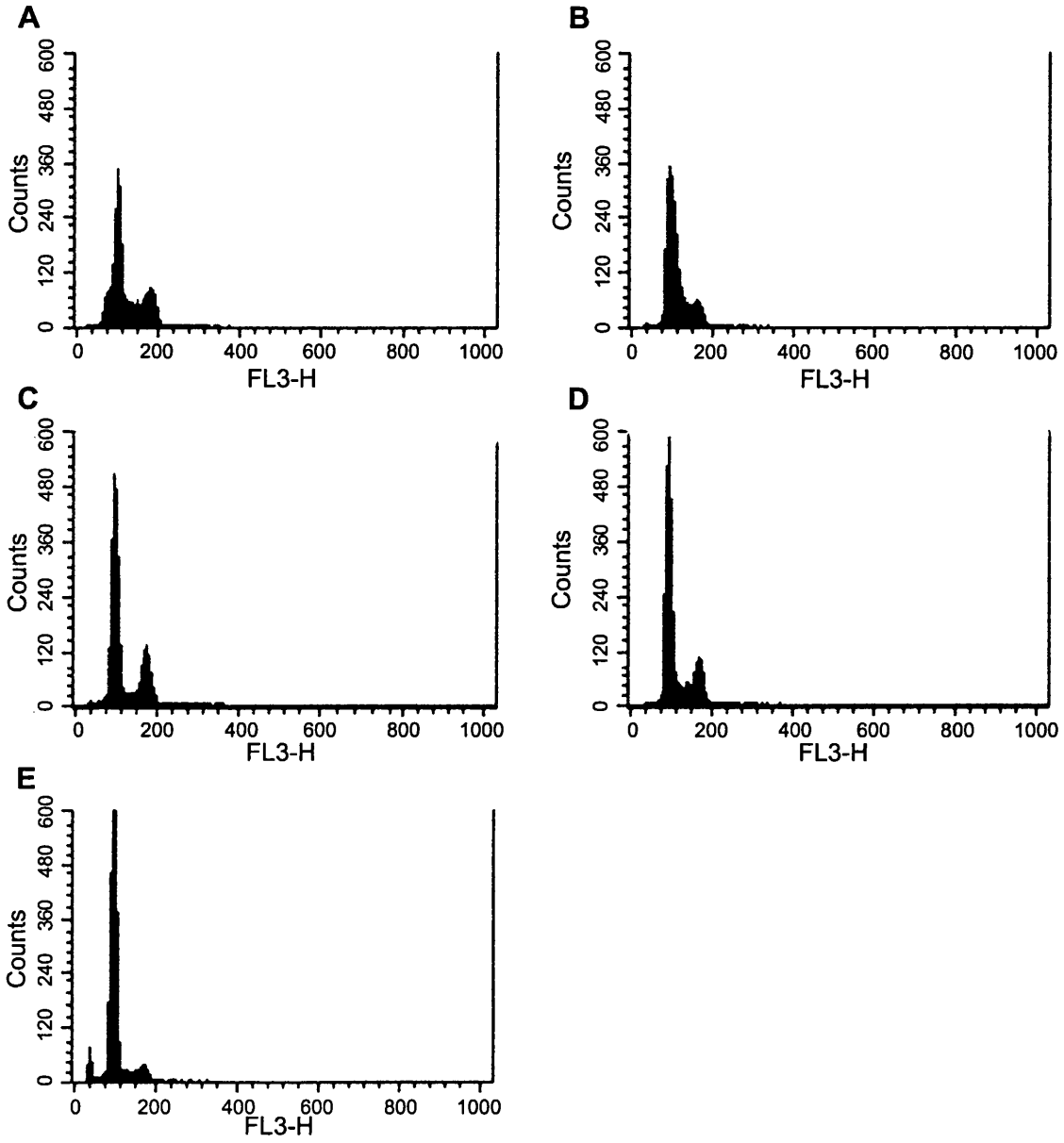


Figure 1. Analysis of Cos-7 cells by FACS flow cytometry. The peak to the left corresponds to cells in the G1 phase of the cell cycle; the peak to the right corresponds to cells in the G2 phase. The shoulder between the 2 peaks corresponds to the S phase. Different cell cycles show variable features, according to the treatment described in Patients, Materials, and Methods. Asynchronously growing cells were given a characteristic trimodel profile *(A);* the thymidine-treated cells *(B)* with 84% of cells were in the G0/1 phase. The release period in which cells were thymidine free *(C)* is shown, as are aphidicolin-treated cells 12 h *(D)* and 30 h *(E)* after administration of aphidicolin. A total of 88% and 94% of Cos-7 cells arrested in the G1-S phase of the cell cycle after incubation with aphidicolin for 12 h and 30 h, respectively.

Distribution of Mutations in the Core Gene B Cell Epitopes and the C-Terminus Region

Overall, 34 mutations were distributed in different known B cell epitope regions within the core gene (table 1). Twenty-one mutations (62%) occurred in residues 74–89, whereas 3 mutations (8%) occurred in residues 107–118 and 9 mutations (30%) occurred in residues 128–135 (table 4). Of 21 mutations in the first B cell epitope (aa residues 74–89), 8 mutations occurred at aa position 80. Thirteen (59%) of 22 C-terminus mutations occurred within aa residues 148–160 (table 4).

Cellular Distribution of Core Protein

Expression of HBcAg was predominantly nuclear in 13 samples, both nuclear and cytoplasmic in 3 samples , and predominantly cytoplasmic in 10 samples. Of the 13 samples that showed predominantly nuclear expression of HBcAg, 9 were HBeAg

Table 4. Correlation between sequence and hepatitis B core antigen (HBcAg) localization.

Patient code	HBcAg localization	Clinical status[a]	T helper 1			20	21							49	T helper 50											69
			12	13			26	27	29	35	38	40			51	55	56	57	58	59	63	64	66	67	68	
			T	V			S	V	D	S	Y	E			H	L	R	Q	A	I	G	E	M	T	L	
1A[1]	N																									
1A[2]	C		S	M								D	T		N									F		
1B	C	Remission	S									D								V	D		I	F		
2A	N																									
2B	C	Remission																H	V			D				
3A	N																									
3B	C	Remission																				D	I		S	
4A	N									A																
4B	C	Active	S									D														
5A	N																					D	I		S	
5B	C	Active																				D	I		S	
6A	N																									
6B	C	Active																				D				
7A	NC		S											T												
7B	NC	Remission	S						A	E																
8A	N							T		Q	F	D					I									
8B	C	Active						T		L	F	D					I							N		
9A	N											D														
9B	NC	Active	S	L								D								V		D				
10A	C											D								V		D				
10B	N	Active						A				D						H	V	V		D				
11A	N											D					I									
11B	C	Active										D					I									
12A	N																									
13A	N																									
14A	N																									

NOTE. B cell and T helper epitopes and C-terminus areas, with their boundaries and wild-type variants. Amino acids are denoted by a single-letter code and are numbered from the beginning of HBcAg. Only positions at which changes occurred are shown, so the relative proportion of epitopic to nonepitopic areas is skewed in favor of regions where substitutions occurred. C, predominantly cytoplasmic distribution; N, predominantly nuclear distribution; NC, mixed nuclear and cytoplasmic localization.

[a] Status of 11 samples obtained from patients who were anti-hepatitis B e positive during a later phase of the study, according to whether the samples were obtained during the remission phase or the active-disease phase.

positive and 4 were anti-HBe positive (i.e., samples 8A, 9A, 11A, and 10B; (note that 3 of the samples that showed nuclear expression were initial samples) (see below). Only 1 patient whose sample (sample 1A[1]) showed a predominantly cytoplasmic distribution of HBcAg was HBeAg positive; in other words, 9 patients were anti-HBe positive. In keeping with this finding, none of the patients whose samples had both nuclear and cytoplasmic distribution of HBcAg were HBeAg positive; in other words, all of these patients were anti-HBe positive. Also, we observed a form of nuclear expression associated with the membrane, by use of standard microscopy (results not shown). Figure 2 and tables 2 and 4 detail the patterns of core localization in 3 patient-derived HBV core genes cloned into pRK5 and transfected into G0/1-arrested Cos-7 cells. Consequently, the first finding of the present study is that cytoplasmic distribution of HBcAg is closely associated with HBeAg negativity/anti-HBe positivity.

Core gene sequences from the initial samples showed nuclear distribution. Tables 2 and 4 show the results of core gene substitutions in 11 initial samples (either HBeAg-positive or anti-HBe–positive samples) and in 3 HBeAg-positive samples obtained at a single point in time. Twelve of these 14 samples shown predominantly nuclear distribution (only 1 of the samples obtained later during the anti-HBe–positive phase [sample 10B] had nuclear localization; a total of 13 samples had predominantly nuclear localization). Eight of 13 samples (all but 1 of which were from HBeAg-positive patients) did not have any mutation. The remaining 5 samples (3 of which were anti-HBe positive) had mutations in either the C-terminal region and/or the B cell epitopes. Two of these samples (samples 9A and 5A) had mutations only in B cell epitopes (A74V/G [2 possible variants for genotype D] and T80A, respectively). An additional 2 samples contained mutations in both regions; sample 6A showed mutations T80A, Q130P, C153G, and T155S, and sample 11A showed mutations S135P, Q151R, and T155S. Only 1 sample (sample 8A) with a C-terminus mutation alone (C153G) was observed. As noted below, a second set of samples that were obtained from the same 5 patients later during the anti-HBe–positive phase developed further mutations and shifted to the cytoplasm. Finally, HBcAg from the remaining 2

Table 4. *(Continued.)*

		B cell													B cell					B cell						C-terminus									
70	73	74	77	79	80	81	87	89	90	92	93	96	97	106	107	109	113	118	119	127	128	130	131	135	136	150	151	153	155	156	160	177	179	180	181	183
T	G	V/G	E	P	A	S	S			N	M	K				T	E			R		P	A	P			R	G	S	P	T	Q	R	E	S	
	S					A																														
	S					A	N						I				D															P	A		
																																	P	A		
																															A					
					T																													P		
			Q	Q	P																		P											P		
					T																															
					T																											K				
					T																	Q						C	T							
	V				T					V												Q						C	T							
					Q											M																		P		
					Q															L		Q						C	T							
		V																										C								
		G		Q	I									E										T						L						
		A																																		
		A									V						Q					Q												P		
		A																					P		F											
																								S			Q		T							
					V		G				T													Q			Q		T							

initial samples (samples 10A and 7A) showed either predominantly cytoplasmic distribution or both nuclear and cytoplasmic distribution, respectively (figure 2, *sample 10A*). Therefore, all but 2 of the initial samples that were obtained showed predominantly nuclear expression. It was unusual to observe C-terminus and/or B cell epitope mutations and nuclear localization of samples that were HBeAg positive.

Accumulation of B cell epitope/C-terminus mutations and HBcAg localization. For the 5 patients with initial samples that contained B cell epitope and/or C-terminus mutations and predominantly nuclear localization (samples 5A, 6A, 8A, 9A, and 11A) (table 2), there was a shift, over time, both in cellular distribution and sequence. Samples that were obtained later during the anti-HBe–positive phase (samples 5B, 6B, 8B, 9B, and 11B) (table 2) had predominantly cytoplasmic distribution (see below): all sequences in the latter group contained additional mutations in the same and/or other regions (table 4). Thus, it appears that there is an association between accumulation of additional mutations and a shift to cytoplasmic distribution.

A shift in cellular distribution to the cytoplasm is linked to substitutions in the carboxy terminal and/or B cell epitopes. Of the 10 samples with predominantly cytoplasmic distribution that were obtained later in the study, 8 contained both C-terminus and B cell epitope mutations (tables 2 and 4). Sample 2B contained only a C-terminus mutation (A160T), whereas sample 10A only had mutations in 2 B cell epitopes (A74V/G and P131A). Sample 10B, although obtained later during the anti-HBe–positive phase, shifted back to the nucleus: in keeping with this finding, all the mutations in sample 10A (the first sample obtained) reverted back to the wild type (i.e., there were no mutations) in sample 10B (figure 2 and table 4). Sample 8A, which was obtained during the initial phase and which had a C-terminus mutation (C153G) (see above), reverted back, and an L156P mutation arose during the anti-HBe–positive phase (table 4, *sample 8B*). An additional aa variation (V74G) occurred in the sample obtained during the later phase (sample 8B). Both of these variants in aa position 74 are specific for genotype D (M.S.J. et al., in press). Both the samples that showed nuclear and cytoplasmic distribution (samples 7B and 9B) contained mutations in both the C-terminus and the B cell epitope regions (table 4). In sequences with mutations, there were no differences in the distribution of those mutations between the 2 clinical groups (i.e., patients with seroconversion from HBeAg to anti-HBe and patients who were continuously anti-HBe positive). Thus, all samples obtained during the later phase showed predominantly cytoplasmic distribution (including both nuclear and cytoplasmic distribution) and contained B cell epitope and/or C-terminus mutations. All but 1 of the samples were anti-HBe positive. As a result, cytoplasmic dis-

q42

q43

q44

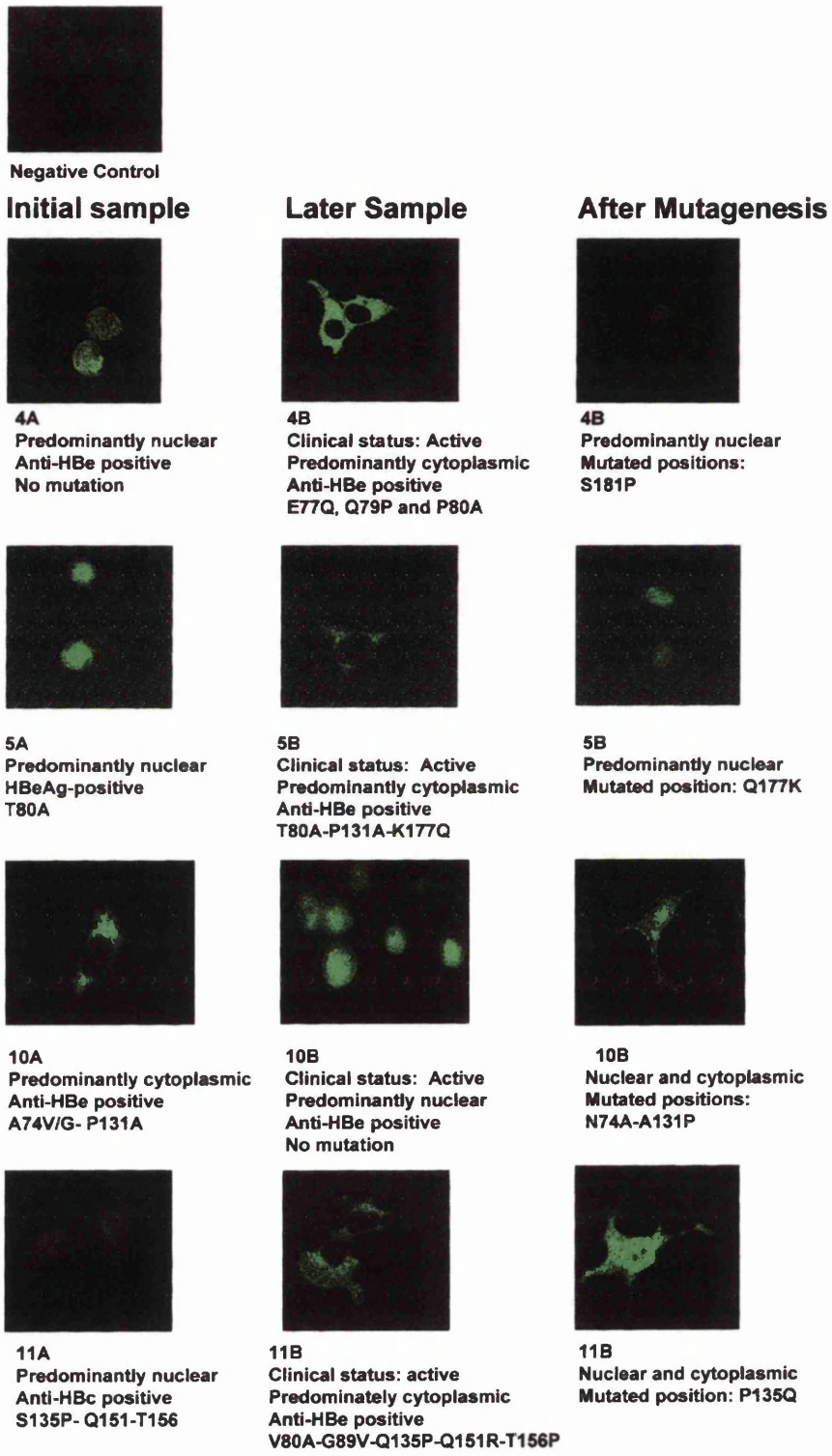


Figure 2. Intracellular localization of hepatitis B core antigen (HBcAg), by use of confocal microscopy. Immunofluorescence staining of HBcAg in Cos-7 cells transfected with constructs that contained clones of initial samples (samples 4A, 5A, 10A, and 11A), and "later" samples (i.e., samples obtained later during the anti–hepatitis B e [HBe]—positive phase) (samples 4B, 5B, 10B, and 11B), and mutated sequences in the B cell epitope and C-terminus of patients with chronic hepatitis B virus infection, as described in the text. Cos-7 cells showed various patterns of predominantly nuclear distribution, predominantly cytoplasmic distribution, and both nuclear and cytoplasmic distribution. Note that only B cell epitope and C-terminus mutations are indicated as variations. The clinical status of later samples is shown. No cells showed predominantly cytoplasmic localization after mutagenesis, indicating that sequence is critical to intracellular distribution. HBeAg, hepatitis B e antigen.

tribution of HBcAg was strongly correlated with the presence of mutations in B cell epitopes and/or the C-terminus region of the C gene as well as with anti-HBe positive status.

Correlation of HBcAg Localization with Clinical Outcome

Of the 11 samples obtained during the later phase, 4 were obtained from patients whose disease was in clinical remission and 7 were obtained from patients who had active disease. Samples from the 4 patients whose disease was in clinical remission showed a shift in the distribution of HBcAg from the nucleus to the cytoplasm after seroconversion to anti-HBe. All but 1 of samples (sample 10B) from the 7 patients who had active disease showed a shift from nuclear to cytoplasmic expression, in a comparison of samples obtained during the earlier and later phases (table 2). Of interest, with the exception of sample 10B, all samples obtained during the active-disease phase and the remission phase contained B cell epitope and/or C-terminus mutations. Consequently, in 10 of 11 samples obtained later during the study, either from patients with active disease or from patients whose disease was in remission, there was a shift from predominantly nuclear to predominantly cytoplasmic distribution, a finding that indicates that there was no clear difference between different clinical outcomes and HBcAg localization.

Cellular Distribution of the Core Protein is Dependent on the Sequence

Some of the shifter samples (table 2) were chosen for site-directed mutagenesis (table 3). Samples 3B, 4B, and 5B had double mutagenesis, at the same time, on both the C-terminus and B cell epitopes. Because sample 10A did not contain a C-terminus mutation but contained mutations in 2 different B cell epitopes (table 4), it was selected for double mutagenesis on these epitopes. All experiments were designed to revert the mutated amino acid to the expected amino acid variant for each genotype, which was also that observed in the first sample of the pair obtained (or the second sample, as in the case of patient 10). All clones were sequenced to confirm reversion. After the sequence in the C-terminal and/or B cell epitopes of the later sample was reverted back to the original sequence, a shift back to the original nuclear distribution and/or both nuclear and cytoplasmic distribution was observed. After single amino acids in B cell epitopes underwent mutation, 4 of 5 samples shifted back partially (i.e., they localized to both the nucleus and the cytoplasm) (figure 2, *sample 11B*). In contrast, mutagenesis of C-terminus mutations resulted in predominantly nuclear localization in all cases (figure 2, *samples 4B and 5B*; table 3). In keeping with these findings, all double-mutated samples shifted back to either nuclear distribution (results not shown) or nuclear and cytoplasmic distribution (figure 2, *sample 10B*). Consequently, reversion of C-terminus mutations leads to shifting of HBcAg back to the nucleus. Reversion of B cell epitope mutations leads to a mixed nuclear and cytoplasmic distribution. The effect of dual mutations is the same as that of C-terminus mutations.

Precore Variants

The precore stop codon mutation was found at aa position 28 q48
of the precore region in 8 of 13 samples with predominantly cytoplasmic localization; it was found in 3 of 13 samples with predominantly nuclear distribution. Consequently, the precore region did not appear to influence the effect of C gene localization. Note that the pre-C region was not within the cloned product (table 2), so it did not directly affect localization.

DISCUSSION

The entire HBV genome is extremely stable in the early high-replicative HBeAg-positive phase, when the immune response is considered to be weak [29]. However, after seroconversion of HBeAg to anti-HBe, multiple clusters of mutations appear in different regions of HBcAg—in particular, in immunodom- q49
inant epitopes (table 1). Several studies have shown that the amino-acid variability of HBcAg is related to the severity of liver disease, both in the HBeAg and the anti-HBe phases [7–8, 10–11, 28, 30]. Previous studies have reported HBcAg distribution in the cytoplasm, nucleus, or both, of infected or transfected cells [17–20, 22–24, 26, 30–34], but none of the studies evaluated naturally mutated sequences. Recently, Kawai et al. [35] found that hepatocytes that exhibited cytoplasmic expression of HBcAg contained more core promoter double mutations (T1762 and A1764) than did hepatocytes that exhibited nuclear expression. Studies of liver biopsy specimens have suggested that the intracellular distribution of HBcAg correlates with disease activity, with cytoplasmic expression being associated with more severe disease [17, 19, 25, 34, 36]. The cell membrane of HBV-infected hepatocytes may also express HBcAg during chronic infection [1, 37]. However, whether intracellular distribution of HBcAg is associated with liver damage is unclear.

Previous studies have evaluated biopsy specimens for which the additional influence of host factors has made it difficult to delineate the mechanisms responsible for intracellular distribution of HBcAg. The present study has confirmed previous observations that there is a close association between the topographic distribution of HBcAg in vitro and the inflammatory status of the liver in vivo (figure 2 and table 2). For the first time, it has been shown that the distribution of expressed HBcAg is strongly associated with the viral sequence (tables 2 and 4). The presence of cytoplasmic expression of HBcAg correlated with the existence of mutations in the B cell epitopes and/or the carboxy terminus of the protein. These experiments

were performed in the presence of aphidicolin, to block cells in the G1/S phase; consequently, any shift to cytoplasmic expression of HBcAg was not a result of dissolution of the nuclear membrane. However, to prove that sequence was critical in HBcAg localization, reversion of C-terminus and/or B cell epitope mutations back to the wild type was performed; in most cases, reversion shifted HBcAg localization back to the nucleus (table 3). Some possible explanations for this shifting are presented below.

Machida and et al. [4, 38] reported an antigenic site within the C-terminus region, known as "HBicAg" (hepatitis B inner cAg; aa residues 148–160), to be a potential B cell antigenic determinant [39]. In the present study, 13 of 22 C-terminus mutations occurred within aa residues 148–160 (table 4). If the arginine-rich C-terminus is a B cell epitope, that might explain the mutations that occurred in this region. Because this region is also known to affect nuclear localization, mutations would explain the shift of HBcAg to the cytoplasm. This is in line with observations that escape mutations appeared in the B cell epitopes of patients with ongoing disease [8] (tables 2 and 4), because double mutagenesis of both a B cell epitope and C-terminus residue led samples shifted back to both nuclear and cytoplasmic distribution and/or predominantly nuclear distribution (results not shown).

Although B cell epitope mutations tended to shift HBcAg to the cytoplasm, they was not as effective as C-terminus mutations, which probably direct HBcAg intracellular trafficking by a different mechanism(s). A majority of samples with B cell mutagenesis had localization in both the nucleus and the cytoplasm (figure 2, *sample 11B*), whereas all samples with C-terminus mutations showed predominantly nuclear distribution (figure 2, *samples 4B and 5B*). This was also observed in 2 of 3 samples with double mutations (figure 2, *sample 10B*). The reason underlying this phenomenon is unclear and deserves more investigation.

In the present study, we clearly showed that, in samples with both nuclear and cytoplasmic localization, HBcAg could be associated with the nuclear membrane (results not shown). This unusual pattern has been reported elsewhere [40]. Because there is no sharp boundary between the G1 and S phases, transport of HBcAg to the nuclear membrane may occur during the transition between phases. It may be that nuclear/cytoplasmic localization may be a transition step between the 2 compartments [19]. Alternatively, there may have been high levels of nuclear HBV DNA with inefficient transportation from the nucleus. However, there is no strict division between predominantly nuclear and predominantly cytoplasmic localizations of core protein in vivo, because core particles usually are present in both compartments [30, 26].

For a majority of patients, a predominantly cytoplasmic distribution in vitro was found in samples obtained during the active-disease phase and the remission phase (table 2). These samples (samples $1A^1$, $1A^2$, 2B, and 3B) selected more mutations in Th epitopes than in the C-terminus region and/or B cell epitopes (table 4). Of the samples obtained from 7 patients who had active disease, 6 showed a shift from nuclear to cytoplasmic expression in a comparison of samples obtained in the early and later phases (table 2); of these 6 samples, 4 (samples 4B, 6B, 8B, and 11B) had a higher proportion of mutations in B cell epitopes than in other regions (table 4). All samples with active disease were anti-HBe positive; 4 had predominantly nuclear distribution (see above). In keeping with this finding, all 4 samples that had predominantly nuclear localization and were obtained from patients who were continuously anti-HBe positive were obtained from patients with active disease. The only apparent correlation of HBcAg sequences in samples 8A, 9A, 10B, and 11A with those in samples from other anti-HBe–positive patients with predominantly cytoplasmic distribution was that fewer mutations were found in these samples, compared with later samples with predominantly cytoplasmic expression (table 4). Although this unusual expression has been reported elsewhere [35, 40–41], it is likely that, after seroconversion of HBeAg to anti-HBe, there is coexistence of both antigen and antibody for some time [42], with concurrent persistence of a low concentration of integrated HBV DNA in the liver after the disappearance of HBeAg [40]. Nevertheless, various studies reported continued liver disease activity after certain patients experienced seroconversion to anti-HBe status that might be related to continued viral replication and secretion with the presence of HBcAg in the nuclei. Even in some patients with cirrhosis and severe chronic hepatitis, distribution of HBcAg has been reported to be predominantly nuclear. The reason underlying this pattern is not clear, but it might be related to the specific patient population studied [16].

In summary, we have shown that subcellular expression of HBcAg is related to serologic patterns and that it depends on the HBcAg sequence, even in samples successively obtained from individual patients. C-terminus and/or B cell epitope mutations are involved in the shifting of HBcAg between the nucleus and the cytoplasm. These results were generated by transfection. Because there is no susceptible cell line that can be truly infected by HBV, the biological importance of such variants, with regard to the viral life cycle and cell behavior, are still unclear.

Acknowledgments

We thank R. Clyton, Angela Rinaldi, and Joy Keen for technical assistance. We gratefully acknowledge Roger Everett for helpful comments on cell-cycle blocking.

References

1. Mondelli M, Vergani GM, Alberti A, et al. Specificity of T lymphocyte cytotoxicity to autologous hepatocytes in chronic hepatitis B virus infection: evidence that T cells are directed against HBV core antigen expressed on hepatocytes. J Immunol **1982**; 129:2773–8.
2. Milich DR. Molecular and genetic aspects of the immune responses to hepatitis B viral antigens. Adv Exp Med Biol **1989**; 257:115–33.
3. Milich DR, McLachlan A. The nucleocapsid of hepatitis B virus is both a T-cell–independent and a T-cell–dependent antigen. Science **1986**; 234:1398–401.
4. Machida A, Ohnuma H, Takai E, et al. Antigenic sites on the arginine-rich carboxyl-terminal domain of the capsid protein of hepatitis B virus distinct from hepatitis B core or e antigen. Mol Immunol **1989**; 26: 413–21.
5. Carman WF. Variation in the core and X genes of hepatitis B virus. Intervirology **1995**; 38:75–88.
6. Boner W, Schlicht HJ, Hanrieder K, Holmes EC, Carman WF. Further characterization of 2 types of precore variant hepatitis B virus isolates from Hong Kong. J Infect Dis **1995**; 171:1461–7.
7. Carman WF, Thursz M, Hadziyannis S, et al. Hepatitis B e antigen negative chronic active hepatitis: hepatitis B virus core mutations occur predominantly in known antigenic determinants. J Viral Hepat **1995**;2: 77–84.
8. Carman WF, Boner W, Fattovich G, et al. Hepatitis B virus core protein mutations are concentrated in B cell epitopes in progressive disease and in T helper cell epitopes during clinical remission. J Infect Dis **1997**; 175:1093–100.
9. Okumura A, Ishikawa T, Yoshioka K, Yuasa R, Fukuzawa Y, Kakumu S. Mutation at codon 130 in hepatitis B virus (HBV) core region increases markedly during acute exacerbation of hepatitis in chronic HBV carriers. J Gastroenterol **2001**; 36:103–10.
10. Ehata T, Omata M, Yokosuka O, Hosoda K, Ohto M. Variations in codons 84–101 in the core nucleotide sequence correlate with hepatocellular injury in chronic hepatitis B virus infection. J Clin Invest **1992**; 89:332–8.
11. Chuang WL, Omata M, Ehata T, et al. Precore mutations and core clustering mutations in chronic hepatitis B virus infection. Gastroenterology **1993**; 104:263–71.
12. Khakoo SI, Ling R, Scott I, et al. Cytotoxic T lymphocyte responses and CTL epitope escape mutation in HBsAg, anti-HBe positive individuals. Gut **2000**; 47:137–43.
13. Kakimi K, Guidotti LG, Koezuka Y, Chisari FV. Natural killer T cell activation inhibits hepatitis B virus replication in vivo. J Exp Med **2000**; 192:921–30.
14. Kakimi K, Lane TE, Chisari FV, Guidotti LG. Cutting edge: inhibition of hepatitis B virus replication by activated NK T cells does not require inflammatory cell recruitment to the liver. J Immunol **2001**; 167: 6701–5.
15. Baron JL, Gardiner L, Nishimura S, Shinkai K, Locksley R, Ganem D. Activation of a nonclassical NKT cell subset in a transgenic mouse model of hepatitis B virus infection. Immunity **2002**; 16:583–94.
16. Hadziyannis SJ, Lieberman HM, Karvountzis GG, Shafritz DA. Analysis of liver disease, nuclear HBcAg, viral replication, and hepatitis B virus DNA in liver and serum of HBeAg vs. anti-HBe positive carriers of hepatitis B virus. Hepatology **1983**; 3:656–62.
17. Wu JC, Lee SD, Wang JY, et al. Correlation between hepatic hepatitis B core antigen and serum hepatitis B virus-DNA levels in patients with chronic hepatitis B virus infections in Taiwan. Arch Pathol Lab Med **1987**; 111:181–4.
18. Hsu HC, Su IJ, Lai MY, et al. Biologic and prognostic significance of hepatocyte hepatitis B core antigen expressions in the natural course of chronic hepatitis B virus infection. J Hepatol **1987**; 5:45–50.
19. Akiba T, Nakayama H, Miyazaki Y, Kanno A, Ishii M, Ohori H. Relationship between the replication of hepatitis B virus and the localization of virus nucleocapsid antigen (HBcAg) in hepatocytes. J Gen Virol **1987**; 68:871–7.
20. Chu CM, Yeh CT, Chien RN, Sheen IS, Liaw YF. The degrees of hepatocyte nuclear but not cytoplasmic expression of hepatitis B core antigen reflect the level of viral replication in chronic hepatitis B virus infection. J Clin Microbiol **1997**; 35:102–5.
21. Chu CM, Yeh CT, Sheen IS, Liaw YF. Subcellular localisation of hepatitis B core antigen in relation to hepatocyte regeneration in chronic hepatitis B. Gastroenterology **1995**; 109:1926–32.
22. Park YN, Han KH, Kim KS, Chung JP, Kim S, Park C. Cytoplasmic expression of hepatitis B core antigen in chronic hepatitis B virus infection: role of precore stop mutants. Liver **1999**; 19:199–205.
23. Brechot C, Pourcel C, Louise A, Rain B, Tiollais P. Presence of integrated hepatitis B virus DNA sequences in cellular DNA of human hepatocellular carcinoma. Nature **1980**; 286:533–5.
24. Gowans EJ, Burrell CJ, Jilbert AR, Marmion BP. Patterns of single- and double-stranded hepatitis B virus DNA and viral antigen accumulation in infected liver cells. J Gen Virol **1983**; 64:1229–39.
25. Ballare M, Lavarini C, Brunetto MR, et al. Relationship between the intrahepatic expression of 'e' and 'c' epitopes of the nucleocapsid protein of hepatitis B virus and viraemia. Clin Exp Immunol **1989**; 75: 64–69.
26. Naoumov NV, Antonov KA, Miska S, Bichko V, Williams R, Will H. Differentiation of core gene products of the hepatitis B virus in infected liver tissue using monoclonal antibodies. J Med Virol **1997**; 53:127–38.
27. Yeh CH, Wong SW, Fung Y, Ou JH. Cell cycle regulation of nuclear localization of hepatitis B virus core protein. Proc Nat Acad Sci USA **1993**; 90:6459–63.
28. Carman WF, McIntyre G, Hadziyannis S, Fattovich G, Alberti A, Thomas HC. Core protein evolution after selection of hepatitis B precore mutants and correlation with disease severity. Viral Hepat Liver Dis **1994**; 273–6. q57
29. Hannoun C, Horal P, Lindh M. Long-term mutation rates in the hepatitis B virus genome. J Gen Virol **2000**; 81:75–83.
30. Yamada G, Nakane PK. Hepatitis B core and surface antigens in liver tissue: light and electron microscopic localization by the peroxidase-labeled antibody method. Lab Invest **1977**; 36:649–59.
31. Nakopoulou L, Adraskelas N, Stefanaki K, Zacharoulis D, Hadziyannis S. Expression of HBsAg and HBcAg in liver tissue: correlation with disease activity. Histol Histopathol **1992**; 7:493–9.
32. Fang JW, Wu PC, Lai CL, Lo CK, Meager A, Lau JY. Hepatic expression of interferon-alpha in chronic hepatitis B virus infection. Dig Dis Sci **1994**; 39:2014–21.
33. Bock CT, Schwinn S, Schroder CH, Velhagen I, Zentgraf H. Localization of hepatitis B virus core protein and viral DNA at the nuclear membrane. Virus Genes **1996**; 12:53–63.
34. Serinoz E, Varli M, Erden E, et al. Nuclear localization of hepatitis B core antigen and its relations to liver injury, hepatocyte proliferation, and viral load. J Clin Gastroenterol **2003**; 36:269–72.
35. Kawai K, Horiike N, Michitaka K, Onji M. The effects of hepatitis B virus core promoter mutations on hepatitis B core antigen distribution in hepatocytes as detected by laser-assisted microdissection. J Hepatol **2003**; 38:635–41.
36. Tong SP, Li JS, Vitvitski L, Trepo C. Active hepatitis B virus replication in the presence of anti-HBe is associated with viral variants containing an inactive pre-C region. Virology **1990**; 176:596–603.
37. Trevisan A, Realdi G, Alberti A, Ongaro G, Pornaro E, Meliconi R. Core antigen-specific immunoglobulin G bound to the liver cell membrane in chronic hepatitis B. Gastroenterology **1982**; 82:218–22.
38. Machida A, Ohnuma H, Tsuda F, et al. Phosphorylation in the carboxyl-terminal domain of the capsid protein of hepatitis B virus: evaluation with a monoclonal antibody. J Virol **1991**; 65:6024–30.
39. Sallberg M, Ruden U, Wahren B, Magnius LO. Immune recognition of linear antigenic regions within the hepatitis B pre-C and C-gene translation products using synthetic peptides. J Med Virol **1994**; 42: 7–15.

40. Wu PC, Fang JW, Lai CL, et al. Hepatic expression of hepatitis B virus genome in chronic hepatitis B virus infection. Am J Clin Pathol **1996**; 105:87–95.
41. ter Borg F, ten Kate FJ, Cuypers HT, et al. A survey of liver pathology in needle biopsies from HBsAg and anti-HBe positive individuals. J Clin Pathol **2000**; 53:541–8.
42. Maruyama T, Kuwata S, Koike K, et al. Precore wild-type DNA and immune complexes persist in chronic hepatitis B after seroconversion: no association between genome conversion and seroconversion. Hepatology **1998**; 27:245–53.

HBV genotypes, core gene variability and ethnicity in the Pacific region

Mohammed S Jazayeri[1], Ashraf A Basuni[1], Graham Cooksley[2], Stephen Locarnini[3] and William F Carman[1,4]

1- Institute of Virology, University of Glasgow, Scotland.

2- Clinical Research Centre, Royal Brisbane Hospital Foundation, Brisbane, Australia

3- Research and Molecular Development, Victorian Infectious Diseases Reference Laboratory, Melbourne, Australia.

4- West of Scotland Specialist Virology Centre, Gartnavel General Hospital, Glasgow, Scotland.

*Corresponding author: WF Carman

West of Scotland Specialist Virology Centre, Gartnavel General Hospital.

PO Box 16766, Glasgow, G12 0ZA, Scotland

Tel: +44 (0) 141-2110085

Fax: +44 (0) 141-2110082

E.mail: w.carman@vir.gla.ac.uk

Number of words in the abstract: 115.

Number of words in the text: 2776.

Abstract

Background/Aims: The world-wide distribution of HBV genotypes follows a geographic pattern under the influence of ethnic background. *Methods:* 48 core genes from four Pacific islands were compared with the following findings. *Results:* First, island-specific variant substitutions were found for only two out of four islands. Second, 11 amino acid and 90 nucleotide changes specific for Pacific genotypes C and D were defined. Third, the nucleotide diversity of genotype C (all but one were silent) was greater than that of genotype D. *Conclusions:* These results suggest an early appearance of genotype C in the Pacific with few subsequent amino acid changes because of shared immunological responses across the region followed by random silent changes, some of which reflect isolation of individual island populations. Genotype D appeared later.

Keywords: HBcAg variability, HBV adrq- subtype, Pacific-specific core gene sequences.

Introduction

HBV genomes can be classified into at least eight genotypes (A to H) based on nucleotide divergence between the strains (1-4). Subtypes are based on antigenic typing, but correlate broadly with genotypes. Some subtypes can be found in more than one genotype; hence, these confer additional heterogeneity within the genotypes. Subdeterminants of w (w1 to w4) have allowed the further characterisation of ayw and adw subtypes into a total of 10 subtypes: ayw1, ayw2, ayw3, ayw4, ayr, adw2, adw4, adw4q-, adrq+ and adrq-. The q determinants were found to be absent from adr strains in the Pacific region, thus, defining adr and adw as either q+ or q- (5). Between strains, ayw and adw subtypes demonstrate considerable genetic heterogeneity (2). Genotypes and subtypes probably evolved in ethnic backgrounds over centuries (6) and the relationship of these ten subtypes to genomic groups A to H has been established.

HBV genotypes have a characteristic geographic distribution, largely in agreement with subtype distribution (7). Genotype A is widely distributed in Western Europe, USA and Sub-Sahara Africa as well as Asia as far as east as the Philippines. Genotypes B and C belong to the indigenous population of South East Asia (1, 8). In addition, genotype C is found in the populations of the Pacific Islands, and can be sub-differentiated geographically by subtype in this region. As compared to strains from South East Asia, two adrq- strains formed a distinct cluster within the Pacific area (7). Genotype D has been found worldwide with its highest prevalence in the Mediterranean area, the Middle East, and South Asia (7). Genotype E is found in West and South Africa. Genotype F is found in South and Central America. Genotype G has been found in France and the USA (3). The distribution of genotype H seems to be restricted so far to the Northern part of Latin America including Central America and Mexico (4).

It has been proposed that HBV diverged into genotypes according to the distribution of humans among the different continents (9). Well-known waves of migration have occurred over time: the prevalence of different HBV genotypes reflects the origin of the immigrants and other patterns of migration (10). The presence of genotype D in Central America reflects the westward immigration of Spanish in whom genotypes A and D are common (11). Yamashita et al, (12). analysed the geographical interface of adr (genotype C) and adw (genotypes A and/or B) strains from Japan. They found an apparent south-to-north gradient of the r determinant and concluded that Japan may have been originally inhabited by people infected with adw (Philippines and Indonesia), and later, by people from the Asian continent infected with adr (South East Asia). In keeping, the HBV type in populations living on Pacific Islands correlates with immigration from South East Asia (genotype C) (13-15). An alternative hypothesis, because of the great diversity of genotype C in the Pacific (Basuni et al, unpublished) is that genotype C migrated from the Pacific to South-East Asia.

The hypothesis that the outcome of HBV infection is influenced by HLA-associated immune response genes has been evaluated in many studies (16-20). There is an association between a self-limited course of acute infection and the presence of certain HLA alleles: HLA-DR B1-1302 in West Africa (21); DR B1-1301 and 1302 among Caucasians (22); and DR13A in Asians (23). In keeping, immunodominant CTL epitopes restricted by HLA-A2404 have been identified in East Asia (24). Collectively, these results suggest that the epitopes for CTL/Th recognition might be different on account of the diverse distribution of HLA antigens in different geographic regions. Analysis of patients with different HLA haplotypes revealed that several sequences within the HBV proteins, in particular core, could induce significant levels of T cell response. Consequently, it is of relevance to sequence geographically divergent HBV genomes to describe the extent of natural variation (25).

0. Wu PC, Fang JW, Lai CL, et al. Hepatic expression of hepatitis B virus genome in chronic hepatitis B virus infection. Am J Clin Pathol **1996**; 105:87–95.
1. ter Borg F, ten Kate FJ, Cuypers HT, et al. A survey of liver pathology in needle biopsies from HBsAg and anti-HBe positive individuals. J Clin Pathol **2000**; 53:541–8.
42. Maruyama T, Kuwata S, Koike K, et al. Precore wild-type DNA and immune complexes persist in chronic hepatitis B after seroconversion: no association between genome conversion and seroconversion. Hepatology **1998**; 27:245–53.

HBV genotypes, core gene variability and ethnicity in the Pacific region

Mohammed S Jazayeri[1], Ashraf A Basuni[1], Graham Cooksley[2], Stephen Locarnini[3] and William F Carman[1,4]

1- Institute of Virology, University of Glasgow, Scotland.

2- Clinical Research Centre, Royal Brisbane Hospital Foundation, Brisbane, Australia

3- Research and Molecular Development, Victorian Infectious Diseases Reference Laboratory, Melbourne, Australia.

4- West of Scotland Specialist Virology Centre, Gartnavel General Hospital, Glasgow, Scotland.

*Corresponding author: WF Carman

West of Scotland Specialist Virology Centre, Gartnavel General Hospital.

PO Box 16766, Glasgow, G12 0ZA, Scotland

Tel: +44 (0) 141-2110085

Fax: +44 (0) 141-2110082

E.mail: w.carman@vir.gla.ac.uk

Number of words in the abstract: 115.

Number of words in the text: 2776.

Abstract

Background/Aims: The world-wide distribution of HBV genotypes follows a geographic pattern under the influence of ethnic background. *Methods:* 48 core genes from four Pacific islands were compared with the following findings. *Results:* First, island-specific variant substitutions were found for only two out of four islands. Second, 11 amino acid and 90 nucleotide changes specific for Pacific genotypes C and D were defined. Third, the nucleotide diversity of genotype C (all but one were silent) was greater than that of genotype D. *Conclusions:* These results suggest an early appearance of genotype C in the Pacific with few subsequent amino acid changes because of shared immunological responses across the region followed by random silent changes, some of which reflect isolation of individual island populations. Genotype D appeared later.

Keywords: HBcAg variability, HBV adrq- subtype, Pacific-specific core gene sequences.

Introduction

HBV genomes can be classified into at least eight genotypes (A to H) based on nucleotide divergence between the strains (1-4). Subtypes are based on antigenic typing, but correlate broadly with genotypes. Some subtypes can be found in more than one genotype; hence, these confer additional heterogeneity within the genotypes. Subdeterminants of w (w1 to w4) have allowed the further characterisation of ayw and adw subtypes into a total of 10 subtypes: ayw1, ayw2, ayw3, ayw4, ayr, adw2, adw4, adw4q-, adrq+ and adrq-. The q determinants were found to be absent from adr strains in the Pacific region, thus, defining adr and adw as either q+ or q- (5). Between strains, ayw and adw subtypes demonstrate considerable genetic heterogeneity (2). Genotypes and subtypes probably evolved in ethnic backgrounds over centuries (6) and the relationship of these ten subtypes to genomic groups A to H has been established.

HBV genotypes have a characteristic geographic distribution, largely in agreement with subtype distribution (7). Genotype A is widely distributed in Western Europe, USA and Sub-Sahara Africa as well as Asia as far as east as the Philippines. Genotypes B and C belong to the indigenous population of South East Asia (1, 8). In addition, genotype C is found in the populations of the Pacific Islands, and can be sub-differentiated geographically by subtype in this region. As compared to strains from South East Asia, two adrq- strains formed a distinct cluster within the Pacific area (7). Genotype D has been found worldwide with its highest prevalence in the Mediterranean area, the Middle East, and South Asia (7). Genotype E is found in West and South Africa. Genotype F is found in South and Central America. Genotype G has been found in France and the USA (3). The distribution of genotype H seems to be restricted so far to the Northern part of Latin America including Central America and Mexico (4).

It has been proposed that HBV diverged into genotypes according to the distribution of humans among the different continents (9). Well-known waves of migration have occurred over time: the prevalence of different HBV genotypes reflects the origin of the immigrants and other patterns of migration (10). The presence of genotype D in Central America reflects the westward immigration of Spanish in whom genotypes A and D are common (11). Yamashita et al, (12). analysed the geographical interface of adr (genotype C) and adw (genotypes A and/or B) strains from Japan. They found an apparent south-to-north gradient of the r determinant and concluded that Japan may have been originally inhabited by people infected with adw (Philippines and Indonesia), and later, by people from the Asian continent infected with adr (South East Asia). In keeping, the HBV type in populations living on Pacific Islands correlates with immigration from South East Asia (genotype C) (13-15). An alternative hypothesis, because of the great diversity of genotype C in the Pacific (Basuni et al, unpublished) is that genotype C migrated from the Pacific to South-East Asia.

The hypothesis that the outcome of HBV infection is influenced by HLA-associated immune response genes has been evaluated in many studies (16-20). There is an association between a self-limited course of acute infection and the presence of certain HLA alleles: HLA-DR B1-1302 in West Africa (21); DR B1-1301 and 1302 among Caucasians (22); and DR13A in Asians (23). In keeping, immunodominant CTL epitopes restricted by HLA-A2404 have been identified in East Asia (24). Collectively, these results suggest that the epitopes for CTL/Th recognition might be different on account of the diverse distribution of HLA antigens in different geographic regions. Analysis of patients with different HLA haplotypes revealed that several sequences within the HBV proteins, in particular core, could induce significant levels of T cell response. Consequently, it is of relevance to sequence geographically divergent HBV genomes to describe the extent of natural variation (25).

In a pilot analysis, we showed that specific core gene motifs (amino acid and nucleotide) could be defined for the Pacific regions in both genotypes C and D (Jazayeri et al., in press). Here, a further 48 sequences have been analysed. We have confirmed our previous results and defined further specific variants that are unique for particular islands, genotypes and subtypes.

Materials and Methods

HBsAg positive samples were collected from four Pacific islands (Fig 1, table 1): Kiribati (14), Vanuatu (14), Fiji (15) and Tonga (15). They were chosen from a previous study of surface variability based on known migration patterns (Basuni, PhD thesis, 2001).

DNA Extraction and PCR

DNA extraction and PCR were done as have been described previously (Jazayeri et al., in press). Briefly, DNA was extracted from the serum samples, then, the 1st round PCR reaction was carried out in 50 µl of a mixture containing 5 µl of the extracted DNA, using standard methodology. The complete core gene was amplified using C1 and C4 primers. The region of the surface gene specifying HBV subtype (amino acid positions 122-160) was amplified using S1, S2Na primers (Jazayeri et al., in press) to allow an allocation of the virus to subtype. For the second round PCR, 1 µl of first round PCR products was added to 49 µl of the reaction mixture with the same composition as the first round except that C1 was replaced by C3A, and S1/S2Na, were replaced by S6C/S7D. PCRs were analyzed by electrophoresis in 1% agarose gel, stained by ethidium bromide, and visualized under UV light.

DNA Sequencing

After PCR product purification, direct sequencing of core and surface genes was done as described previously (Jazayeri et al., in press) and results were analysed using Sequence Navigator software (Applied Biosystems).

Sequence Analysis

After allocating a sequence to an HBV genotype by analysis of the S gene, the core gene amino acid/nucleotide variations that were found in a majority (> 50%) of HBcAg sequences from a particular island were recorded and compared with HBcAg sequences obtained from databases. The adrq- subtype is rarely found in Genbank; of eight available, six were submitted as part of our studies (Jazayeri et al., in press). These, and two isolates (HMA and Cha) published by Norder et al., (7), constituted our reference sequences. No Pacific ayw sequence (excluding this study) has been reported in Genbank. Finally, genotype/subtype-specific nucleotide and amino acid variations were identified. Sequences have been submitted to GenBank, numbered from AY269035 to AY269047, from AY269050 to AY269061 and from AY269064 to AY269086. The accession numbers for the previous 10 samples from the pilot study were AF324097 to AF3224106. After construction of simple Neighbour-joining Trees based on core and surface proteins, further phylogenetic trees were constructed using sequences from the database and from a previous study (Jazayeri et al., in press) for comparison.

Phylogenetic Analysis:

Sequences of core and surface genes were aligned using the BioEdit Package version 5.0.9 (26), and a neighbor joining phylogenetic tree constructed using the Treecon Package (27) employing a Kimura distance matrix (28). Associations were tested by bootstrap re-sampling analysis using 100 replicates (29). Branches with a bootstrap value of greater than 70% were deemed well-supported by the data.

Results

Genotype and subtype prevalence in Pacific islands

Table 1 shows the distribution of genotypes and subtypes of strains studied from individual islands. The following preparations were genotype D: in Kiribati, 13 out of 14; in Fiji, 5 out of 15; in Vanuatu, 1 out of 14; and in Tonga, 2 out of 15. The remaining sequences were genotype C. Subtype adrq- was the main subtype in genotype C-occupied islands while ayw2 was the prevalent one in Kiribati.

Genotype/subtype-specific substitutions common to Pacific region

Comparing genotypes C and D, there were 11 aa variations at 10 positions (table 2): S/T12, I/V27, A/S35, D/E40, I/V59, N/T67, S/T74, D/E83, I/V/T91 and F/I97. These variants were not island-specific with the exception of V91 found only in Vanuatu.

It was also possible to identify nucleotide variants for genotypes C and D, subtypes adrq- and ayw2. Overall, 90 genotype/subtype-specific nucleotide substitutions in 45 positions were found (table 3), of which 67 were silent. 23 were missense (genotype specific) in the 10 positions detailed above. Interestingly, a majority of adrq- strains derived from Vanuatu (but not other islands) contained nucleotide variants identical to ayw (genotype D) in 4 positions: 78, 138, 234 and 255.

Substitutions that distinguish specific islands

Island-specific nucleotide variants were identified which increased or decreased in frequency from West to East (from Vanuatu to Fiji and on to Tonga) (figure 1, table 4). Nucleotide changes at positions C_{36}, T_{109}, C_{114}, G_{255} and T_{267} were found in a majority of strains from Tonga. On the other hand, changes including C_{78}, T_{234} A_{246}, A_{255}, G_{271}, C_{354}, C_{366}, A_{454} and G_{510} were predominant in Vanuatu. However, no island-specific nucleotide variants were detected in Fiji and Kiribati. All nucleotide changes that were specific to an island group, except G_{271} (which corresponded to residue V91 in Vanuatu), were silent mutations.

There were six strains from Tonga (3309, 3415, 3417, 3419, 3428 and 3629) and four from Vanuatu (5017, 5022, 5230 and 5072) of genotype C, adrq- subtype, which showed high variability at the nucleotide level compared to other Pacific strains (results not shown). However, in genotype D-studied strains, the only major heterogeneity was seen in isolate 2143 from Kiribati. All these strains originated from different cities in each island with various HBeAg/anti-HBe status.

Comparison with international database

Choosing the correct reference sequence is mandatory, especially in cross-sectional studies, otherwise over- and/or under-estimations can occur. The adrq- subtype is rarely found in Genbank; of eight available, six were submitted as part of our previous study (Jazayeri et al, in press) and two isolates (HMA and Cha), published by Norder et al., (7), constituted our reference sequences. No ayw sequence from Pacific (excluding this study) has been reported in Genbank. We chose one genotype C from a Chinese (A35) and one Indian genotype D (0123), both from our pilot study for comparison between corresponding Pacific strains. For genotype C, there were 2 amino acid and 12 nucleotide variations between our predominant Pacific genotype C sequence and A35 (Chinese reference) (table 5). However, in genotype D, we found 5 amino acids and 30 nucleotides different between our Pacific D and 0123 (table 6).

Phylogenetic Analysis

Genotype C Tree

Figure 2 shows that all the genotype C sequences, from the pilot study (bold codes), database (bold italic codes) and adrq- strains (from this study, italics), grouped into two major clusters. The first is completely occupied by Pacific adrq-, including one of the database sequences, X75656 (7); a majority of strains belonging to Vanuatu composed a major sub-cluster. The

second cluster contains adrq+ sequences (including from the database) derived mainly from S.E Asian populations, which are heterogeneous.

Genotype D Tree

Figure 3 shows alignment of all genotype D sequences. They grouped into 2 major clusters. The first, similar to the genotype C tree, was entirely made up of Pacific strains (italic codes). The second cluster contained two sub-clusters: the first had those from Western populations (a majority from Europe); the second had a combination of Asian and a few European sequences. Inside the Euro-Asian cluster, there was an Indian sub-cluster. Unlike genotype C, the Pacific genotype D shows relative homogeneity with only a few nucleotide mutations, indicating that genotype D arrived recently in the Pacific as a relatively homogenous population. We have been unable to identify a likely some of this migration in this study.

Discussion

The prevalence of hepatitis B in the South Pacific is amongst the highest in the world. This geographic area comprises a unique region of admixing two HBV genotypes, C and D, with specific identities which have not been identified in other regions. In a pilot study, analysis of ten HBV sequences from 4 Pacific islands showed unique changes in both genotypes C and D (Jazayeri et al., in press). This provoked us to investigate another 48 sequences which confirmed those.

The intra- and inter-Pacific island nucleotide variations showed that a majority of such variants, despite being silent, were within the known HBcAg immune epitopes, mainly CD4. Phylogenetic analysis revealed that both genotypes belonged to a separate Pacific cluster (figures 2-3). Comparison with databases samples supported our conclusions. Adrq+ is the prevalent subtype in South-East Asia, while adrq- is only found in Oceania. Geographical transition from Asian adrq+ to South American adw4q- has been suggested to be via the adrq- subtype in Oceania (5). The gradient of nucleotide and aa variations from west to east

in our study (shared sequences between Tonga and Fiji and, on the other hand, variations between Tonga and Vanuatu), are most consistent with the hypothesis of migration of Polynesian people from Southern China through Melanesia and Fiji and their radiation across the Pacific to fill the Polynesian triangle in different times (13, 15, 30-31). In contrast, back migration from Polynesia to Melanesia could also be responsible for this variation gradient (see below). Of interest is that 10 strains from Vanuatu, all genotype C isolates, constitute a subcluster in genotype C tree (figure 2, cluster B). However, two isolates from Vanuatu were extremely divergent and found within subcluster A. On the other hand, within the subcluster A, 6 strains from Tonga constitute a separate branch (see results). The presence of different strains from all islands (in particular Fiji and Tonga) in cluster A may indicate admixture of isolates in Fiji due to its geographical location between other 3 islands (fig 1). The gradient of nucleotide substitutions between 3 islands in table 3 showed commonality between Fiji and Tonga, rather than Vanuatu. This is in keeping with the history of an isolated Vanuatu.

The contrast between the silent mutations that predominate within Pacific genotype C sequences and non-synonymous changes that separate Pacific from the S.E Asian sequences is suggestive of an interaction with human genetic variants, especially HLA, in the evolution of geographically separated HBV lineages. Once in the Pacific, no further immune mediated amino acid selection seems to have occurred, probably because there are only a few Pacific HLA types. On the other hand, the significant nucleotide variability is in keeping with random mutation in isolated island population.

In contrast, genotype D strains including those from databases showed that the Pacific islands contain a unique homogenous cluster separated from other genotype D groups. The few D sequences on other islands show a similar homogeneity, indicating recent contact between Kiribati (which is almost an isolated island in Micronesia with limited transfer of virus-see figure 1) and the genotype C-dominated islands. This suggests that genotype D

arrived in the Pacific as a relatively homogenous population and has been distributed by admixing with stable human populations (Basuni, unpublished data). As Indian immigrants usually have genotype D, one might expect Indians to be the origin of genotype D strains in this region. However, our results did not show such an acquisition of these new strains by migration, there was substantial variability between these two ethnic groups. This is best seen as the two separate clusters in figure 3. The study on S genes by Basuni et al, also revealed no strong evidence for Indian admixture in Pacific genotype D populations. In fact, results from our previous pilot study showed that at the nucleotide level the Pacific genotype D sequences were closer to Caucasian strains rather than Indian (results not shown). Thus, a Pacific origin for genotype D cannot be ruled out.

Using phylogenetic network analysis (Basuni, unpublished data), Pacific sequences can be shown to shared an ancestral lineage most closely related to Japan and China, suggesting an immediate Asian origin for genotype C (while, the Pacific genotype D sequences contained a completely separate cluster from India). The hypothesized migration of South-East Asian peoples out into the Pacific correlates with their commonality of HBV sequences similar to the study by Norder et al (7): 37 adrq- strains in genotype C as well as 21 ayw2 in genotype D formed a separate cluster that belonged only to the Pacific area (figures 2-3). Thus, the divergence of HBV genomes in this area is limited, in keeping with the known low variability between genotypes B and C.

To the best of our knowledge, the HBV C genes of Pacific peoples has never been studied. There may be implications for the design of therapeutic vaccines, but, at a minimum, these studies reveal that both C and S genes of HBV have evolved in similar ways in the Pacific. Overall, the high diversity for genotype C (including silent mutations) and the island-specific sequences, suggest a long history of evolution and isolation for genotype C-islands. In

contrast, the low diversity of genotype D indicates a recent, limited, spread over the rest of the Pacific with a single lineage.

Fig 1- Pacific map showing boundaries of the Pacific islands and surrounding countries. Fiji and Vanuatu located in Melanesia, Tonga and Kiribati located in Polynesia and Micronesia, respectively.

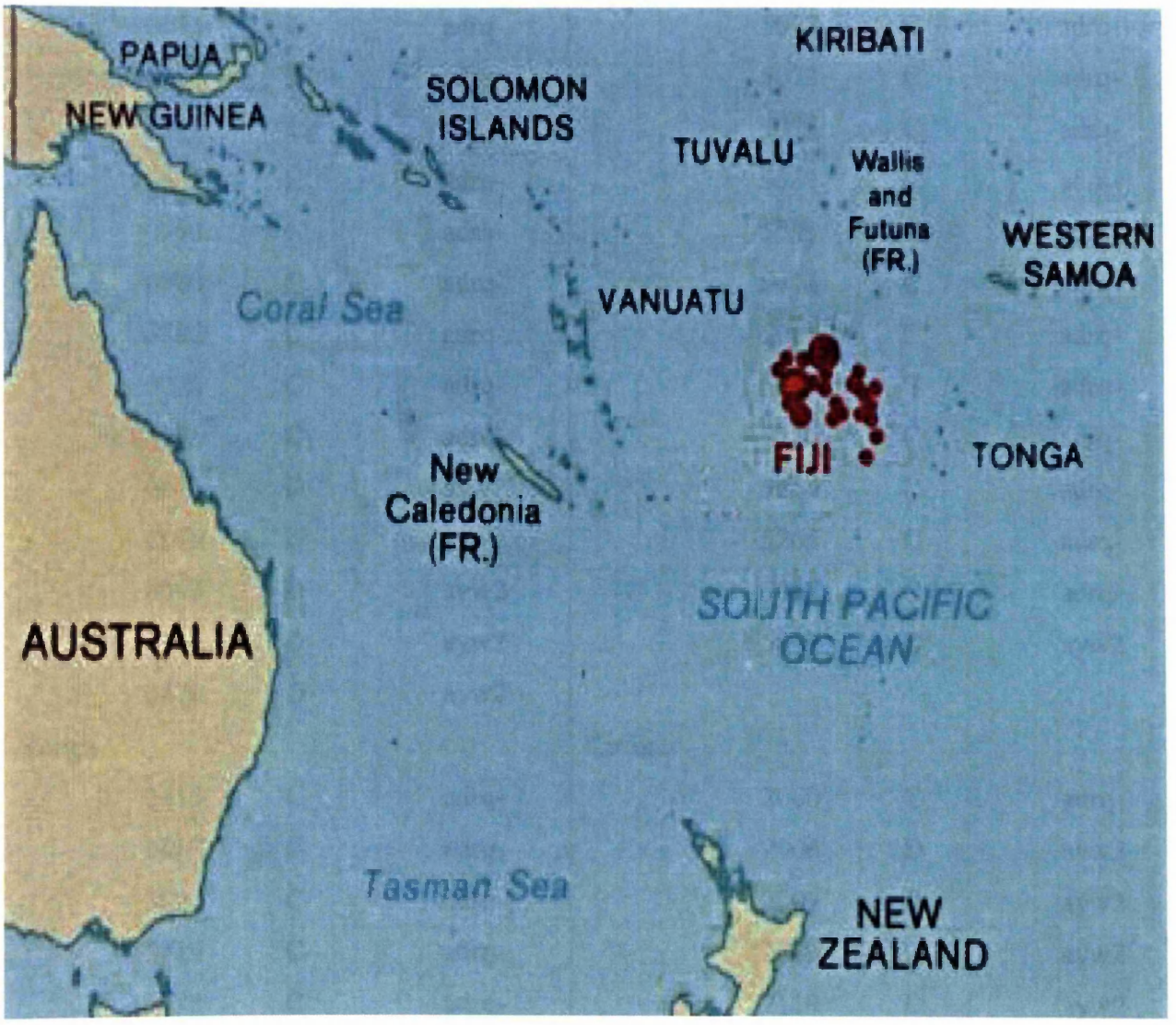


Table 1- Origin and Distribution of HBsAg positive sera that were used in this study as well as genomic/subgenomic identification of Pacific islands strains.

Island	Sample Code	Genotype	Subtype	Island	Sample Code	Genotype	Subtype
Fiji				Vanuatu			
	0142	C	adrq-		5022	C	adrq-
	0143	C	adrq-		5072	C	adrq-
	0314	C	adrq-		5073	C	adrq-
	0349	C	adrq-		5075	C	adrq-
	0480	C	adrq-		5081	C	adrq-
	0550	C	adrq-		5093	C	adrq-
	0800	C	adrq-		5114	C	adrq-
	0880	C	adrq-		5311	C	adrq-
	0026	C	adrq-		5186	C	adrq-
	0075	C	adrq-		5230	C	adrq-
	0078	D	ayw2		5264	C	adrq-
	0079	D	ayw2		5265	C	adrq-
	0092	D	ayw2		5017	C	adrq-
	0456	D	ayw2		5029	D	ayw2
	0470	D	ayw2				
Tonga				Kiribati			
	3415	C	adrq-		2080	C	adrq-
	3417	C	adrq-		2006	D	ayw2
	3097	C	adrq-		2007	D	ayw2
	3419	C	adrq-		2019	D	ayw2
	3428	C	adrq-		2119	D	ayw2
	3629	C	adrq-		2127	D	ayw2
	3099	C	adrq-		2317	D	ayw2
	3221	C	adrq-		2483	D	ayw2
	3365	C	adrq-		2109	D	ayw2
	3369	C	adrq-		2110	D	ayw2
	3519	C	adrq-		2117	D	ayw2
	3309	C	adrq-		2143	D	ayw2
	3509	C	adrq-		2039	D	ayw2
	3414	D	ayw2		2084	D	ayw2
	3343	D	ayw2				

Table 2. The most frequent differences in core gene amino acid sequences between genotypes/subtypes of the Pacific region.

Amino acid Position	Genotype D Subtype ayw	Genotype C Subtype adrq-
12	T	S
27	V	I
35	A	S
40	D	E
59	I	V
67	T	N
74	T	S
83	D	E
91	T	I/V*
97	F	I

Each specific amino acid was found in a majority of strains from each genotype/subtype and the combination was unique for that genotype/subtype. Amino acids are represented by single letters and are numbered from the beginning of the core gene.

V* was only seen in adrq- strains from Vanuatu.

Table 3- Unique C-gene nucleotide differences which permit genotypes C and D classification.

Nucleotide Position	Genotype D Subtype ayw	Genotype C Subtype adrq-	Nucleotide Position	Genotype D Subtype ayw	Genotype C Subtype adrq-
34	A	T	200	C	A
51	G	T	204	A	G
66	C	T	219	T	A
75	T	G	221	C	G
78*	C	T	228	A	G
79	G	A	234‡	T	C
81	A	-	243	T	C
90	T	C	249	C	A
93	A	C	250	C	T
96	T	C	255†	A	G/T
103	G	T	272	C	T
105	G	A	288	G	A
120	T	G	289	T	A
135	T	G	308	T	C
138Φ	G	A	324	C	T
159	T	A	342	G	T
174	A	T	348	A	T
175	A	G	387	C	T
180	T	G	390	A	C
183	C	T	396	T	C
189	A	T	418	C	T
192	A	G	435	G	A
193	C	T			

Note: Nucleotides represented by single letters and numbered from the beginning of the core gene. Dash lines represent no variant specified for that position. Bold numbers and letters indicate amino acid altered positions (missense mutations).

*: A majority of adrq- strains from Vanuatu had C_{78}.

Φ: A majority of adrq- strains from Vanuatu had C_{138}.

‡: adrq- strains from Vanuatu contained T_{234}.

†: A_{255} seen in a majority of samples from Vanuatu despite belonging to adrq- subtype.

Table 4. Nucleotide differences between prevalent strains in Fiji, Tonga and Vanuatu.

Nucleotide Positions/Changes	Vanuatu	Fiji	Tonga
C36T	1	2	6
C78T	9	6	3
T109C	1	2	6
C114T	0	4	12
A138G	3	9	12
T234C	11	0	3
A246G	10	1	1
G255T	0	2	6
A255T	9	6	2
T267C	0	2	6
G271A	11	1	0
C354T	10	2	0
C366T	10	1	0
A454C	7	2	0
G501A	3	6	12
G510A	11	7	2

Table 5- Alignment of nucleotide differences between the Pacific genotype C strains compared with a Chinese genotype C.

Sample Code	Country	15	66	75	78	114	138	175	195	207	204	246	255	261	270	271	354	366	390	501	510
A35	China	G	C	T	T	T	A	A	G	C	C	G	A	C	T	G	T	T	T	G	G
0142	Fiji	T	T	G	T	T	A	G	A	T	C	G	T	C	C	A	T	T	A	A	A
0143	Fiji	T	T	G	T	T	A	G	A	T	C	G	T	C	C	A	T	T	C	A	G
0314	Fiji	T	T	G	T	C	A	G	A	T	C	G	T	T	C	A	T	T	C	G	A
0349	Fiji	T	T	G	T	T	A	G	A	T	C	G	T	C	C	A	T	T	C	A	A
0480	Fiji	T	T	G	T	T	A	G	A	T	C	G	T	T	C	A	T	T	C	G	A
0550	Fiji	A	T	G	T	C	A	G	A	T	C	G	G	T	C	A	T	T	C	G	A
0800	Fiji	T	T	G	T	T	A	G	A	T	C	G	T	T	C	A	C	T	C	G	A
0880	Fiji	T	T	G	C	T	G	G	A	T	C	A	A	T	C	G	C	C	C	A	G
0078	Fiji	T	T	G	T	C	A	G	A	T	C	G	G	T	C	A	T	T	C	G	A
0075	Fiji	T	T	G	T	C	A	G	A	T	C	G	T	T	C	A	T	T	C	G	A
2080	Kiribati	T	T	G	C	T	G	G	A	T	T	G	T	T	C	G	T	T	C	A	G

3415	Tonga	T	T	G	T	C	A	G	A	T	C	G	T	T	C	A	T	T	C	G	A
3417	Tonga	T	T	G	T	C	A	G	A	T	C	G	T	T	C	A	T	T	C	G	A
3097	Tonga	T	T	G	T	C	A	G	A	T	C	G	T	T	T	A	T	T	C	G	A
3419	Tonga	T	T	G	T	C	A	G	A	T	C	G	T	T	C	A	T	T	C	G	A
3428	Tonga	T	T	G	T	C	A	G	A	T	C	G	T	T	C	A	T	T	C	G	A
3629	Tonga	T	T	G	T	C	G	G	A	T	C	G	T	T	C	A	T	T	C	G	C
3099	Tonga	T	T	G	C	C	A	G	A	T	C	G	G	T	C	A	T	T	C	G	A
3221	Tonga	T	T	G	T	T	A	G	A	T	C	G	T	T	C	A	T	T	C	G	A
3365	Tonga	T	T	G	T	C	A	G	A	T	C	G	G	T	C	A	T	T	C	G	A
3369	Tonga	T	T	G	T	C	A	G	A	T	C	G	G	T	C	A	T	T	C	G	A
3519	Tonga	T	T	G	T	C	A	G	A	T	C	G	G	T	C	A	T	T	C	G	A
3309	Tonga	T	T	G	T	C	A	G	A	T	C	G	G	T	C	A	T	T	C	G	A
3509	Tonga	T	T	G	T	C	A	G	A	T	T	A	G	C	C	A	T	T	C	A	A
5022	Vanuatu	T	T	G	C	T	G	G	A	T	T	A	A	T	C	G	C	C	C	A	G
5072	Vanuatu	T	T	G	T	T	A	G	A	T	C	G	A	T	C	G	T	T	C	G	G
5073	Vanuatu	T	T	G	C	T	G	G	A	T	T	A	A	T	C	G	C	C	C	A	G

5075	Vanuatu	T	T	G	C	T	G	G	A	T	T	A	T	T	C	G	C	C	C	A	G
5081	Vanuatu	T	T	G	C	T	G	G	A	T	T	A	T	T	C	G	C	C	C	G	A
5093	Vanuatu	T	T	G	C	T	G	G	A	T	T	A	A	T	C	G	C	C	C	A	G
5114	Vanuatu	T	T	G	T	T	G	G	A	T	T	A	A	T	C	G	C	C	C	A	G
5311	Vanuatu	T	T	G	C	T	G	G	A	T	T	A	T	T	C	G	C	C	C	A	G
5186	Vanuatu	T	T	G	C	T	G	G	A	T	T	A	A	T	C	G	C	C	C	A	G
5230	Vanuatu	T	T	G	T	T	A	G	A	T	C	G	T	T	C	A	T	T	C	A	A
5264	Vanuatu	T	T	G	C	T	G	G	A	T	T	A	A	T	C	G	C	C	C	A	G
5265	Vanuatu	T	T	G	C	T	G	G	A	T	T	A	A	T	C	G	C	C	C	A	G
5017	Vanuatu	T	T	A	T	T	A	G	A	T	C	G	T	T	C	A	T	T	C	G	A

Table 6- HBcAg nucleotide differences between Pacific genotype D and Indian samples.

Nucleotide Position	Pacific	Indian
12	C	T
55	T	C
57	G	T
75	T	A
78	C	A
103	G	T
105	G	A
111	G	A
120	T	A
132	T	G
138	G	A
189	A	G
207	T	C
220	A	G
221	C	G
228	A	G
238	G	A
239	C	T
243	T	C
255	A	G
267	C	T
288	G	A

291	C	T
294	A	G
298	T	C
345	T	C
346	C	A
366	T	C
387	C	T
418	C	T
501	A	G

Fig 2. Neighbour joining phylogenetic tree of core gene sequences of genotype C from Pacific adrq- strains (italic), adr subtypes from the pilot study (bold) and database (bold italic). Sequences belonged to Pacific island shown by first island name letter after their sample code; F, Fiji; K, Kiribati; T, Tonga and V for Vanuatu. Strains rooted with sample I98. Figure shows bootstrap value of ≥70% and scale denotes percent diversity. Viet= Vietnam. Sequence variations of Pacific strains grouped into two sub-clusters, A and B.

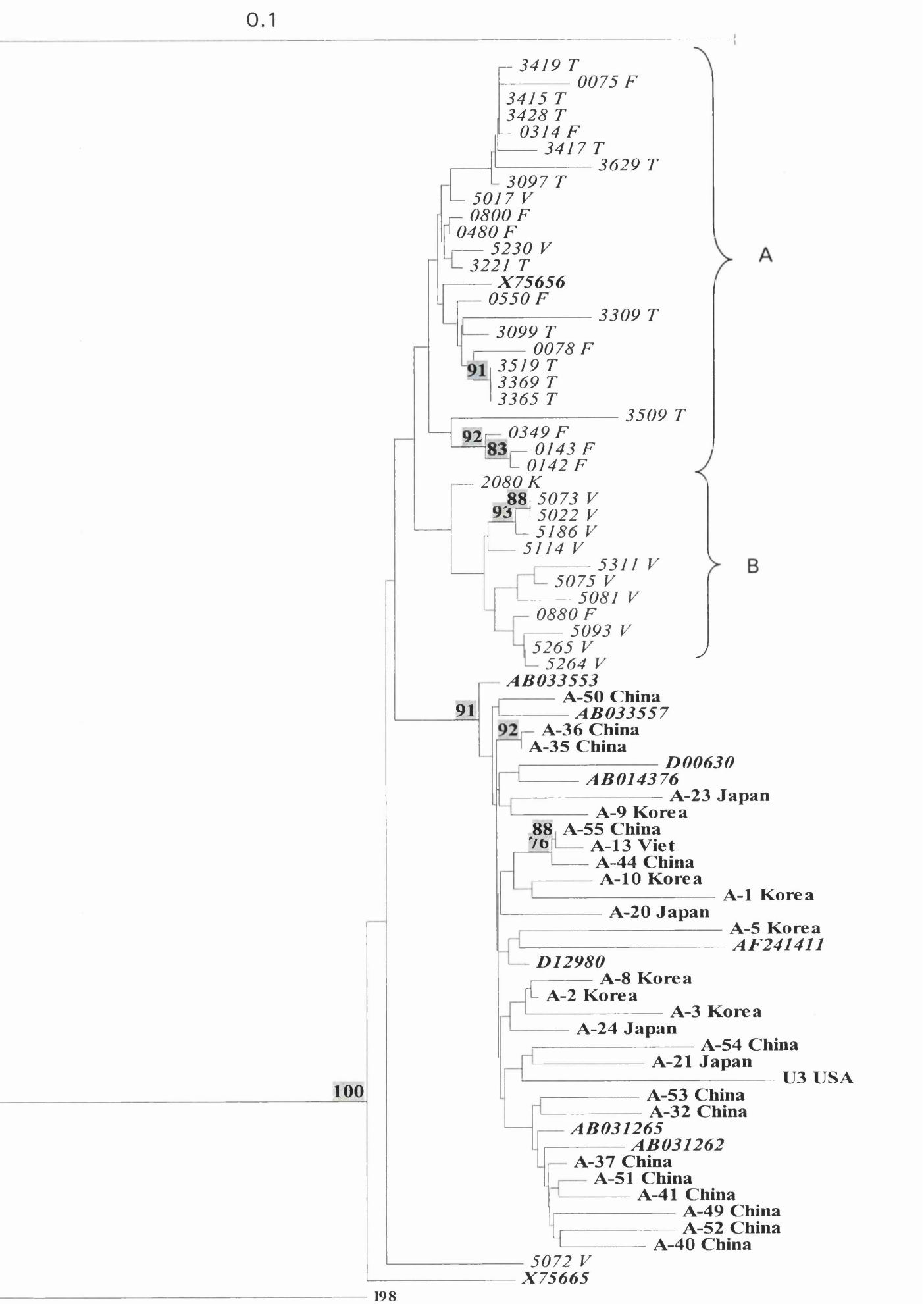
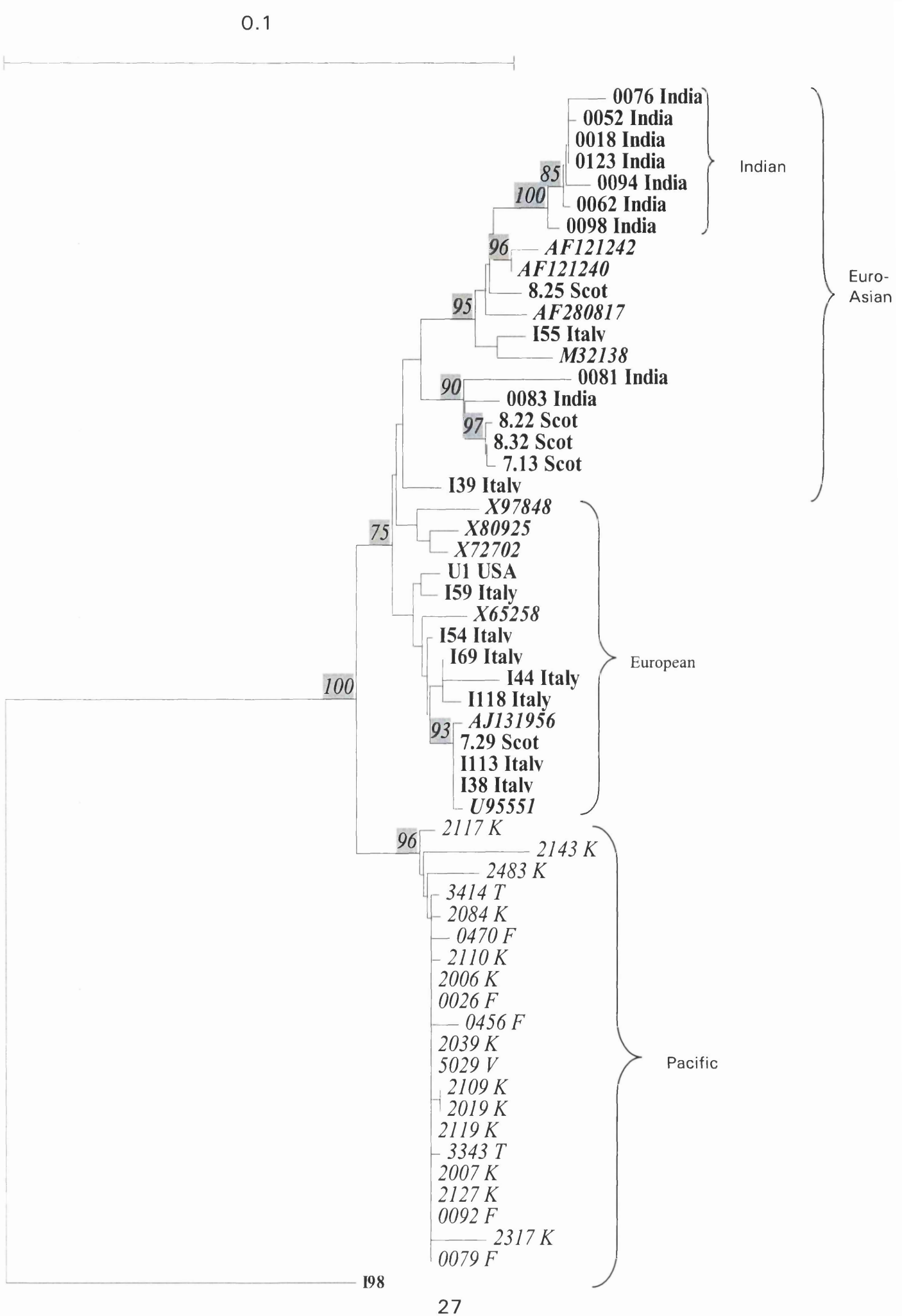

0.1
3419 T
0075 F
3415 T
3428 T
0314 F
3417 T
3629 T
3097 T
5017 V
0800 F
0480 F
5230 V
3221 T
X75656
0550 F
3309 T
3099 T
0078 F
91
3519 T
3369 T
3365 T
3509 T
92
0349 F
83
0143 F
0142 F
A
2080 K
88
5073 V
93
5022 V
5186 V
5114 V
5311 V
5075 V
5081 V
0880 F
5093 V
5265 V
5264 V
B
AB033553
A-50 China
91
AB033557
92
A-36 China
A-35 China
D00630
AB014376
A-23 Japan
A-9 Korea
88
A-55 China
76
A-13 Viet
A-44 China
A-10 Korea
A-1 Korea
A-20 Japan
A-5 Korea
AF241411
D12980
A-8 Korea
A-2 Korea
A-3 Korea
A-24 Japan
A-54 China
A-21 Japan
U3 USA
100
A-53 China
A-32 China
AB031265
AB031262
A-37 China
A-51 China
A-41 China
A-49 China
A-52 China
A-40 China
5072 V
X75665
198

Fig 3. Neighbour joining phylogenetic tree of core gene sequences of genotype D from the Pacific (italic), pilot study (bold) and database (bold italic) ayw strains, rooted with samples I98. Sequences belonged to Pacific island shown by first island name letter after their sample code; F, Fiji; K, Kiribati; T, Tonga and V for Vanuatu. Strains rooted with sample I98. Figure shows bootstrap value of ≥70% and scale denotes percent diversity. Indi=India, Scot= Scotland.


0.1
0076 India
0052 India
0018 India
0123 India
85
0094 India
100
0062 India
0098 India
Indian
96
AF121242
AF121240
8.25 Scot
Euro-
Asian
95
AF280817
I55 Italy
M32138
0081 India
90
0083 India
97
8.22 Scot
8.32 Scot
7.13 Scot
I39 Italy
X97848
75
X80925
X72702
U1 USA
I59 Italy
X65258
I54 Italy
I69 Italy
European
100
I44 Italy
I118 Italy
AJ131956
93
7.29 Scot
I113 Italy
I38 Italy
U95551
2117 K
96
2143 K
2483 K
3414 T
2084 K
0470 F
2110 K
2006 K
0026 F
0456 F
2039 K
5029 V
Pacific
2109 K
2019 K
2119 K
3343 T
2007 K
2127 K
0092 F
2317 K
0079 F
I98

References:

1-Okamoto, H.,Tsuda, F.,Sakugawa, H.,et al (1988). Typing hepatitis B virus by homology in nucleotide sequence: comparison of surface antigen subtypes. J Gen Virol 69, 2575-2583.

2-Norder, H.,Hammas, B.,Lofdahl, S.,Courouce, A. M. & Magnius, L. O. (1992a). Comparison of the amino acid sequences of nine different serotypes of hepatitis B surface antigen and genomic classification of the corresponding hepatitis B virus strains. J Gen Virol 73, 1201-1208.

3- Stuyver, L.,De Gendt, S.,Van Geyt, C.,Zoulim, F.,Fried, M.,Schinazi, R. F. et al. (2000). A new genotype of hepatitis B virus: complete genome and phylogenetic relatedness. J Gen Virol 81, 67-74.

4- Arauz-Ruiz, P.,Norder, H.,Robertson, B. H. & Magnius, L. O. (2002). Genotype H: a new Amerindian genotype of hepatitis B virus revealed in Central America. J Gen Virol 83, 2059-2073.

5- Courouce-Pauty, A. M.,Lemaire, J. M. & Roux, J. F. (1978). New hepatitis B surface antigen subtypes inside the ad category. Vox Sang 35, 304-308.

6- Esteban, J., Martell M, Carman WF, Gomez JEsteban, J., Martell M, Gomez J (1999). The impact of rapid evolution of the hepatitis viruses. In Origin and evolution of viruses, pp. 345-365. Edited by E. Domingo, Webster RG, Holland JI. London: Academic press.

7- Norder, H.,Courouce, A. M. & Magnius, L. O. (1993). Complete nucleotide sequences of six hepatitis B viral genomes encoding the surface antigen subtypes ayw4, adw4q-, and adrq- and their phylogenetic classification. Arch Virol Suppl 8, 189-199.

8- Kidd-Ljunggren, K.,Ekdahl, K.,Oberg, M.,Kurathong, S. & Lolekha, S. (1995). Hepatitis B virus strains in Thailand: genomic variants in chronic carriers. J Med Virol 47, 454-461.

9- Telenta, P. F.,Poggio, G. P.,Lopez, J. L.,Gonzalez, J.,Lemberg, A. & Campos, R. H. (1997). Increased prevalence of genotype F hepatitis B virus isolates in Buenos Aires, Argentina. J Clin Microbiol 35, 1873-1875.

10- Kidd-Ljunggren, K.,Miyakawa, Y. & Kidd, A. H. (2002). Genetic variability in hepatitis B viruses. J Gen Virol 83, 1267-1280.

11- Arauz-Ruiz, P.,Norder, H.,Visona, K. A. & Magnius, L. O. (1997). Molecular epidemiology of hepatitis B virus in Central America reflected in the genetic variability of the small S gene. J Infect Dis 176, 851-858.

12- Yamishita, Y.,Kurashina, S.,Miyakawa, Y. & Mayumi, M. (1975). South-to-north gradient in distribution of the r determinant of hepatitis B surface antigen in Japan. J Infect Dis 131, 567-569.

13- Diamond, J. M. (1988). Expression train to Polynesia. Nature 336, 307-308.

14-Lum, J. k, Rickards, O., Ching, C and Cann, R. L. (1994). Polynesian mitochondrial DNAs reveal three deep maternal lineage clusters.Hum Biol 66, 567-90.

15- Hagelberg, E. & Clegg, J. B. (1993). Genetic polymorphisms in prehistoric Pacific islanders determined by analysis of ancient bone DNA. Proc R Soc Lond B Biol Sci 252, 163-170.

16- Bottcher, B.,Wynne, S. A. & Crowther, R. A. (1997). Determination of the fold of the core protein of hepatitis B virus by electron cryomicroscopy. Nature 386, 88-91.

17- Hillis, W. D.,Hillis, A.,Bias, W. B. & Walker, W. G. (1977). Associations of hepatitis B surface antigenemia with HLA locus B specificities. N Engl J Med 296, 1310-1314.

18- Sampliner, R. E.,Bias, W. B.,Carney, E.,Hillis, A. & Hillis, W. D. (1981). HLA antigens and HBV infection: evaluation in the chronic carrier state and in a large family. Tissue Antigens 18, 247-251.

19- Thio, C. L.,Carrington, M.,Marti, D.,O'Brien, S. J.,Vlahov, D.,Nelson, K. E., et al. (1999). Class II HLA alleles and hepatitis B virus persistence in African Americans. J Infect Dis 179, 1004-1006.

20-Bhimma R, Hammond MG, Coovadia HM, Adhikari M, Connolly CA. HLA class I and II in black children with hepatitis B virus-associated membranous nephropathy. Kidney Int. 2002 Apr; 61(4):1510-5

21- Thursz, M. R.,Kwiatkowski, D.,Allsopp, C. E.,Greenwood, B. M.,Thomas, H. C. & Hill, A. V. (1995). Association between an MHC class II allele and clearance of hepatitis B virus in the Gambia. N Engl J Med 332, 1065-1069.

22- Hohler, T.,Gerken, G.,Notghi, A.,Lubjuhn, R.,Taheri, H.,Protzer, et al. (1997). HLA-DRB1*1301 and *1302 protect against chronic hepatitis B. J Hepatol 26, 503-507.

23- Ahn, S. H.,Han, K. H.,Park, J. Y.,Lee, C. K.,Kang, S. W.,Chon, et al. (2000). Association between hepatitis B virus infection and HLA-DR type in Korea. Hepatology 31, 1371-1373.

24- Sobao, Y.,Sugi, K.,Tomiyama, H.,Saito, S.,Fujiyama, S.,Morimoto, et al. (2001). Identification of hepatitis B virus-specific CTL epitopes presented by HLA-A*2402, the most common HLA class I allele in East Asia. J Hepatol 34, 922-929.

25- Norder, H.,Courouce, A. M. & Magnius, L. O. (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.

26-Hall, T.A.. BioEdit: a user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT. Nucl. Acids. Symp. Ser. 1999; 41:95-98.

27-Van de Peer, Y., De Wachter, R. TREECON for Windows: a software package for the construction and drawing of evolutionary trees for the Microsoft Windows environment. Comput. Applic. Biosci. 1994; 10:569-570.

28-Kimura M. A simple method for estimating evolutionary rates of base substitutions through comparative studies of nucleotide sequences. J. Mol. Evolution. 1980; 16:111-120.

29-Felenstein J. Confidence limits on phylogenies: an approach using the bootstrap. Evolution 1985; 39:783-91.

30- Redd, A. J.,Takezaki, N.,Sherry, S. T.,McGarvey, S. T.,Sofro, A. S. & Stoneking, M. (1995). Evolutionary history of the COII/tRNALys intergenic 9 base pair deletion in human mitochondrial DNAs from the Pacific. Mol Biol Evol 12, 604-615.

31- Kayser, M.,Brauer, S.,Weiss, G.,Underhill, P. A.,Roewer, L.,Schiefenhovel, W. & Stoneking, M. (2000). Melanesian origin of Polynesian Y chromosomes. Curr Biol 10, 1237-1246.