Aus dem Bereich Biologie der Medizinischen Fakultät der Universität des Saarlandes, Homburg/Saar und dem Institut für Immunologie/Centre de Recherche Public de la Santé, Luxemburg

Die weltweite Ausbreitung und Evolution von Hepatitis B Virus Genotypen, Subgenotypen und Rekombinanten

---

The global spread and evolution of hepatitis B virus genotypes, subgenotypes and recombinants

DISSERTATION

zur Erlangung des Grades eines Doktors der Naturwissenschaften
der Medizinischen Fakultät
der UNIVERSITÄT DES SAARLANDES

2013

vorgelegt von

Iris Elke Andernach

geb. am: 18. November 1980 in Koblenz

Table of Contents

Index of Figures

Index of Tables

List of Abbreviations

°C	Degrees Celsius
A	Adenine
AA	Amino acid
Ab	Antibody
Ag	Antigen
ALT	Alanine aminotransferase
anti-HBc	Antibody against HBc
anti-HBe	Antibody against HBeAg
anti-HBs	Antibody against HBsAg
BF	Bayes Factor
BLAST	Basic local alignment search tool
bp	Base pair
C	Cytosine
cccDNA	Covalently closed circular DNA
cDNA	Complementary DNA
ChHBV	Chimpanzee hepatitis B virus
cm	Centimetre
CRP	Centre de Recherche Public
ddNTP	Dideoxynucleotide triphosphate
DHBV	Duck hepatitis B virus
DMSO	Dimethylsulfoxide
DNA	Deoxyribonucleic acid
dNTP	Deoxynucleotide triphosphate
DR	Direct repeat
DTT	Dithiothreitol
EDTA	Ethylene diamine tetra acetic acid
ELISA	Enzyme-linked immunosorbent assay
Enh1	Enhancer 1
Enh2	Enhancer 2

ER	Endoplasmic reticulum
FL	Full length sequence
G	Guanine
GAPDH	Glyceraldehyde – 3 – phosphate dehydrogenase
GiHBV	Gibbon hepatitis B virus
HBc	Hepatitis B core protein
HBeAg	Hepatitis B envelope antigen
HBsAg	Hepatitis B surface antigen
HBV	Hepatitis B virus
HBx	Hepatitis B virus x protein
HCC	Hepatocellular carcinoma
HCV	Hepatitis C virus
HDV	Hepatitis Delta virus
HDAg	Hepatitis Delta virus antigen
HDV-Ab	Hepatitis Delta virus antibody
HIV	Human immunodeficiency virus
IgG	Immunoglobulin G
IgM	Immunoglobulin M
IU	International Unit
kb	Kilobases
l	Litre
LB	Luria Bertani Broth
LNS	Laboratoire National de Santé
mg	Milligram
$MgCl_2$	Magnesium chloride
min	Minute
ml	Millilitre
MJN	Median-joining network
MRCA	Most recent common ancestor
mRNA	Messenger RNA
N	Number
NA	Nucleic acid
NAT	Nucleic acid testing
NCBI	National Center for Biotechnology Information
ng	Nanogram
nm	Nanometre
no.	Number
nt	Nucleotide
NUC	Nucleotide/nucleoside analogue
n.a.	Not applicable
OBI	Occult hepatitis B virus infection
OD	Optical Density
ORF	Open reading frame
OuHBV	Orangutan hepatitis B virus

P	HBV polymerase gene
PBS	Phosphate buffered saline
PCR	Polymerase chain reaction
pgRNA	Pregenomic RNA
PHH	Primary human hepatocytes
rcDNA	Relaxed circular DNA
Ref.	Reference
RFLP	Restriction fragment length polymorphism
RNA	Ribonucleic acid
RNase	Ribonuclease
rpm	Rotations per minute
s	Second
S	HBV surface gene
SNP	Single nucleotide polymorphism
sec	Second
s/s/y	Nucleotide substitutions per site and year
T	Thymine
t	Time
Temp	Temperature
tMRCA	Time of evolution from the most recent common ancestor
WHO	World Health Organisation
WHV	Woodchuck hepatitis virus
ε	Encapsidation signal
X	HBV X gene
µg	Microgram
µl	Microliter
µM	Micromolar

Table 1: Amino acid codes

Single letter code	Amino acid name	Abbreviation
A	Alanine	Ala
R	Arginine	Arg
N	Asparagine	Asn
D	Aspartic acid	Asp
C	Cysteine	Cys
E	Glutamic acid	Glu
Q	Glutamine	Gln
G	Glycine	Gly
H	Histidine	His
I	Isoleucine	Ile
L	Leucine	Leu
K	Lysine	Lys
M	Methionine	Met
F	Phenylalanine	Phe
P	Proline	Pro
S	Serine	Ser
T	Threonine	Thr
W	Tryptophan	Trp
Y	Tyrosine	Tyr
V	Valine	Val

Zusammenfassung

Trotz effizienter Impfungen stellt Hepatitis B Virus (HBV) immer noch ein weltweites Gesundheitsproblem dar. Insbesondere in struktur- und entwicklungsschwachen Regionen, wie zum Beispiel in Subsahara-Afrika oder Südost-Asien, ist HBV hoch prävalent. In diesen Regionen ist die Übertragung während der Geburt oder in früher Kindheit der wichtigste Infektionsweg für HBV und führt häufig zu chronischer HBV Infektion.

Auf Grund genetischer Unterschiede wird HBV in acht klassifizierte Genotypen A-H und einen möglichen Genotyp I eingeteilt. Zusätzlich dazu wurde Genotyp J vorgeschlagen. Mit Ausnahme von E, G, H und J, werden die Genotypen in Subgenotypen unterteilt. In Gebieten mit hoher HBV Prävalenz in denen verschiedene Genotypen und Subgenotypen zirkulieren, sind Infektionen mit mehreren Genotypen und Rekombinationen zwischen Genotypen häufig.

Im ersten Teil der vorliegenden Dissertation wird Genotyp E in Afrika charakterisiert. Dieser Genotyp ist in weiten Teilen West-Afrikas verbreitet und ist hoch prävalent, jedoch wurden identische virale Sequenzen tausende Kilometer entfernt voneinander gefunden. Wir zeigen, dass Genotypen E trotz seiner weiten Verbreitung und Dominanz in weiten Teilen West-Afrikas eine überraschend niedrige genetische Diversität hat. Dies deutet darauf hin, dass dieser Genotyp nur eine kurze Evolutionszeit hat. Im Gegensatz dazu ist Genotyp A, der in dieser Region in der Minderheit ist, hoch divers und hat eine lange Evolution durchlaufen. Die weite Ausbreitung, hohe Prävalenz und niedrige genetische Diversität des Genotyp E ist ein derzeit ungelöstes Rätsel, das wir hier in mehreren Studien untersuchen.

Die Ausbreitung von Pathogenen ist eng mit der ihrer natürlichen Wirte vernetzt. Daher ist anzunehmen, dass auch afrikanische Sklaven, die zwischen dem 17. und 19. Jahrhundert von Afrika nach Süd- und Mittel-Amerika deportiert wurden, die afrikanischen HBV Varianten in der Neuen Welt verbreitet hätten. Wir zeigen jedoch, dass Genotyp E dort nur sporadisch auftritt und bestätigen somit das rezente Erscheinen dieses Genotyps. Auch in Haiti, wo mehr als 90% der Bevölkerung Nachfahren afrikanischer Sklaven sind, wurden mehr als 40% der HBV Infektionen durch Subgenotyp A1 verursacht, der sonst hauptsächlich in Ost-Afrika zu finden ist. Weitere 20% der Infektionen sind auf den seltenen Subgenotyp A5 zurückzuführen, der zuvor nur in der Bucht von Benin, einem Dreh- und Angelpunkt des transatlantischen Sklavenhandels, gefunden wurde. Unsere Ergebnisse zeigen weiterhin, dass sich die haitischen HBV A Subgenotypen früh unabhängig von den afrikanischen weiterentwickelt haben, während die heute in West-Afrika dominanten Sub-/Genotypen (E und A3) in Haiti selten sind. Diese Ergebnisse untermauern, dass die dominanten Sub-/Genotypen des heutigen West-Afrikas sich nur nach dem Ende des transatlantischen Sklavenhandels ausgebreitet haben können, und zwar innerhalb der letzten 100 bis 200 Jahre. Dies erklärt auch die niedrige genetische Diversität von Genotyp E. Die hohe Prävalenz dieses Genotyps in weiten Teilen Afrikas lässt weiterhin darauf schließen, dass die HBV Hyperendemie ein relativ neuer Zustand ist und vermutlich auf kontaminierte Injektionsnadeln in wohlgemeinten Massen-Injektionskampagnen der französischen und belgischen Kolonialmächte zurückzuführen ist.

Des Weiteren zeigen unsere phylogenetischen Analysen, dass Genotyp E eine niedrigere genetische Diversität in den südlichen Ländern des Ausbreitungsgebietes besitzt als in den nördlichen. Dies deutet darauf hin, dass Genotyp E zuerst im Norden aufgetreten ist und von dort in die südlichen Länder verbreitet wurde. Wir bestätigen dies durch phylogeografische Analysen, die auf einen Ursprung von Genotyp E in der Gegend von Nigeria hindeuten, von wo aus sich dieser dann rasant durch das westliche Afrika südlich der Sahara ausbreitete.

Weiterhin finden wir eine Substitutionsrate von 1.9×10^{-4} Substitutionen pro Nukleotid und Jahr und demonstrieren, dass Genotyp E erst kürzlich und sehr wahrscheinlich vor 130 Jahren verbreitet wurde. Auch modellieren unsere Analysen einen starken Anstieg der Genotyp E Infektionszahlen und untermauern so das gehäufte rezente Auftreten. Interessanterweise scheinen die Genotyp E Varianten in Haiti das Ergebnis

mehrerer Einschleppungen in der zweiten Hälfte des 20. Jahrhunderts zu sein, was unsere Hypothese unterstützt, dass Genotyp E zur Zeit des Sklavenhandels noch nicht hochprävalent war. Im Ganzen zeigen unsere Ergebnisse, dass die Hyperendemie von HBV und Genotyp E im heutigen Afrika ein überraschend rezentes Phänomen ist und sehr wahrscheinlich das Ergebnis dramatischer Änderungen der viralen Übertragungswege.

In einer weiteren Studie in Subsahara-Afrika charakterisieren wir Hepatitis D Virus (HDV), ein Satelliten-Virus von Hepatitis B Virus, anhand von Proben aus Burkina Faso, der Zentralafrikanischen Republik, dem Tschad und Nigeria. Wir demonstrieren, dass bis zu 50% der HBV Infizierten, nachgewiesen durch Detektion des HBsAg, auch mit Hepatitis D Virus infiziert sind. Die Prävalenzen variierten jedoch in den verschiedenen Kohorten und Ländern. So hatten Kinder aus Burkina Faso mit 20,5% eine überraschend hohe HDV Prävalenz, die fast 10-mal höher war als die ihrer Mütter. Diese Kinder unterliegen einem hohen Risiko, schwere fulminante Hepatitis oder chronische Hepatitis zu entwickeln. Weiterhin waren 50% der Leberpatienten aus der Zentralafrikanischen Republik HDV Antikörper positiv, im Gegensatz zu 1,3% und 6,8% nigerianischer Leberpatienten. Dies deutet darauf hin, dass HDV Superinfektion in der Zentralafrikanischen Republik eine wichtige Ursache chronischer Hepatitis mit möglicherweise schwerem Verlauf ist.

In der untersuchten geografischen Region konnten wir von den acht charakterisierten viralen Typen (sogenannten clades), Typ 1, 5, und 6 detektieren, wobei jedoch HDV-1 dominierte. Interessanterweise waren mehr als 50% der genotypisierbaren HDV und HBV Paare mit HDV-1 und HBV/E infiziert. Während unsere Analysen darauf hinweisen, dass HDV seinen Ursprung in Afrika hat, zeigen wir, dass eine enge Koevolution von HDV-1 und HBV/E unwahrscheinlich ist.

In weiteren Studien charakterisieren wir HBV in Serumproben aus Laos. Dort sind circa 9% der Bevölkerung chronische Träger von HBsAg, dem primären Marker einer Infektion mit Hepatitis B Virus. Weiterhin ko-zirkulieren in Laos verschiedene Subgenotypen der Genotypen B, C und I.

Wir zeigen, dass fast 6% der HBsAg positiven abgelehnten Blutspender mit einer Vielzahl an Sub-/Genotypen infiziert sind. Weiterhin erbringen wir den Nachweis, dass circa 65% der HBV misch-infizierten virale Varianten mit Rekombinationen im Oberflächengen in sich tragen. Diese hatten überraschend diverse Rekombinations-

Stellen und beinhalteten die vorherrschenden Genotypen B und C. Wir demonstrieren, dass die Rekombinanten weitestgehend unterschiedlich sind und sowohl zwischen Spendern als auch innerhalb einzelner Personen variieren. Auch deuten unsere Analysen darauf hin, dass rekombinante virale Varianten zu Lebzeiten der Spender evolviert sind. Weiterhin schlagen wir auf Basis der Analyse viraler Quasispezies eine Substitutionsrate von $1.58x10^{-4}$ Substitutionen pro Nukleotid und Jahr für HBV in Laos vor. Zusammengenommen zeigen unsere Ergebnisse, dass, zumindest in Laos, ein großer Anteil der Hepatitis B Virus Misch-Infektionen zu Rekombinationen führt, die jedoch noch nicht zu neuen dominanten viralen Varianten in der Bevölkerung geführt haben.

Abschließend demonstrieren wir in serologischen und phylogenetischen Analysen in zufällig ausgewählten Blutspendern, dass 45,5% positiv für mindesten einen Hepatitis B Virus Marker sind. Dies zeigt, dass circa die Hälfte der laotischen Bevölkerung in ihrem Leben in Kontakt mit HBV war. Weiterhin erbringen wir den Nachweis, dass die in Laos verwendeten Screening-Methoden überdacht werden sollten. Wie unsere Ergebnisse zeigen, birgt die Analyse von Blutspenden die nur auf HBsAg limitiert ist, ein beträchtliches Restrisiko durch HBsAg-negative okkulte Infektionen in mindestens 3,9% aller Blutspender. Diese sind potentiell DNA positiv und können Hepatitis B Virus übertragen.

Abstract

Despite efficient vaccines, hepatitis B virus (HBV) continues to be a global public health problem. Especially in low income regions, such as sub-Saharan Africa or South-East Asia, HBV is highly endemic. In these regions perinatal or early childhood transmission is considered the most important route of infection for HBV, frequently leading to chronic HBV carriage.

Based on the genetic variability of the DNA genome, hepatitis B virus has evolved into at least eight classified genotypes A–H and a tentative genotype I. Recently a tenth genotype J has been proposed. With the exception of genotypes E, G, H and J, HBV genotypes are divided into subgenotypes with more or less distinct geographic distributions. In highly endemic regions where different HBV genotypes and subgenotypes co-circulate, mixed infections and recombinations are frequent.

In the first parts of the present dissertation, genotype E is further characterized. This genotype has an extensive spread throughout a vast West African crescent where it is highly prevalent. Nevertheless, identical sequences have been found several thousand kilometres apart. We show that genotype E has a surprisingly low genetic diversity, despite its wide spread and dominance in most countries in sub-Saharan Africa. These results indicate an only short time of evolution for genotype E, while genotype A, the minority genotype in this region, is highly divers and has a long time of evolution. Due to its wide spread, high prevalence and low genetic diversity, genotype E represents an unsolved puzzle that we investigated in multiple studies.

As the spread of pathogens is closely linked to their natural hosts, African slaves that were force-migrated to the Americas from the 17th to the early 19th century during

the transatlantic slave trade would have disseminated African HBV strains in the New World. Nevertheless, we show that genotype E is only sporadically found in the Americas, corroborating an only recent emergence of this genotype. We furthermore demonstrate that in Haiti, where >90% of the population are descendants of African slaves, more than 40% of hepatitis B virus infections were caused by subgenotype A1, which today is found mainly in eastern Africa. Another 20% belong to a rare subgenotype, A5, which has been found only in the former Bight of Benin, a former primary slave trading post. Haitian A subgenotypes appear to have separated early from the African subgenotypes, while the most prevalent genotype and subgenotype in West Africa today (E and A3, respectively) are rare in Haiti. These results provide evidence that the dominant sub/-genotypes in West-Africa emerged in the general population only after the end of the transatlantic slave trade, within the last 100 to 200 years, and explain the low genetic diversity of genotype E. The high prevalence of genotype E in much of Africa further suggests that HBV hyperendemicity is a recent phenomenon, probably resulting from well-intended injection mass campaigns using unsafe needles, performed by the French and Belgian colonial powers.

Our phylogenetic analyses also provide evidence that genotype E may have emerged in the Northern part of the genotype E crescent before spreading to the South, as the Southern countries of the genotype E crescent exhibit lower genetic distances than the Northern countries. We confirm this by phylogeographic analyses that suggest an origin of genotype E in the area of Nigeria, before rapidly spreading throughout sub-Saharan Africa.

We find a substitution rate of 1.9×10^{-4} substitutions per site and year and demonstrate that genotype E emerged only recently, most likely within the last 130 years. Together with the strong increase in the effective number of genotype E infections over time, our results corroborate the recent introduction in sub-Saharan Africa. Interestingly, genotype E strains found in Haiti seem to be the result of multiple introductions only in the second half of the 20th century, corroborating an absence of a significant number of genotype E strains in West Africa when and where the slaves were rounded up. Taken together, the results provide strong evidence that the hyperendemicity of HBV and genotype E in today`s Africa is a surprisingly recent phenomenon and likely the result of dramatic changes in the routes of viral transmission.

In an additional study in sub-Saharan Africa we characterize hepatitis delta virus (HDV), a satellite virus of hepatitis B, in samples from Burkina Faso, the Central African Republic, Chad and Nigeria. We demonstrate that up to 50% of hepatitis B surface antigen positive carriers were infected with HDV, with prevalences varying between cohorts and countries. Children from Burkina Faso had a surprisingly high prevalence of HDV, with 20.5% being infected. This prevalence is almost 10-times higher than the one observed in their mothers and these children are at high risk to develop severe fulminant or chronic hepatitis. We furthermore found 50% of liver patients in the Central African Republic to be HDV-antibody positive, indicating that in the Central African Republic, in contrast to Nigeria with only 1.3% and 6.8% prevalences in liver patients, HDV super-infection might be an important cause of chronic hepatitis with possibly frequent severe liver conditions.

Of the eight characterized hepatitis delta clades, we detect clades 1, 5 and 6 in this region, with a predominance of HDV-1. Interestingly, more than 50% of genotypable HDV and HBV pairs revealed both HDV-1 and HBV/E. However, while our results indicate that HDV might have originated in Africa, we demonstrate that a close co-evolution of HDV clade 1 and HBV genotype E seems unlikely.

In additional studies, we analysed sera from Lao PDR, were about 9% of the population are chronic carriers of HBsAg, the primary marker of hepatitis B virus infection. Additionally, multiple subgenotypes of genotypes B, C and I co-circulate in this country.

We demonstrate that about 6% of HBsAg positive rejected blood donors are infected with multiple sub-/genotypes. We furthermore provide evidence, that about 65% of the mixed infected donors showed recombinations in the surface gene that had surprisingly diverse recombination breakpoints and involved the predominant genotypes B and C. We show that recombinant virus strains were largely distinct and varied between donors and within individuals. In addition, our analyses indicate that recombinant strains largely evolved within the lifetime of the donor. In addition, based on the analyses of viral quasipecies, we propose a substitution rate of 1.58×10^{-4} substitutions per site and year for strains in Lao PDR. Thus, our results provide evidence that, at least in Laos, hepatitis B virus mixed infections lead to frequent recombinations that so far have not led to new dominant strains in the population.

Finally, in serologic and phylogenetic analyses on sera from randomly selected blood donors in Lao PDR, we show that 45.5% of the blood donors were positive for at least one of the hepatitis B virus markers. Our results indicate that about half of the Laotian population has been in contact with this virus at one point during their lives. We furthermore provide evidence, that blood screening procedures in Lao PDR need to be reconsidered. We show that the screening of donations solely for HBsAg leaves a sizable risk of hepatitis B virus transmission by HBsAg-negative occult infected donors in at least 3.9% of blood donations. These are potentially DNA positive and could transmit hepatitis B virus.

Chapter I: Introduction

1 Hepatitis B virus

Hepatitis B virus (HBV) had first been described in an Australian Aborigine in 1967 (35). A specific serum antigen had been observed and designated the Australia antigen, which is now known as the hepatitis B surface antigen (HBsAg). This antigen was linked in 1970 to the presence of the so called Dane particle (70), which was later found to be the complete infectious HBV virion (Figure 1) (131, 244, 245).

Today, HBV is the prototype of the virus family *Hepadnaviridae*. Due to similarities in the DNA sequence and the genome organization, Hepadnaviruses are grouped in the genera *Orthohepadnavirus,* with viruses infecting mammals, and *Avihepadnavirus*, infecting birds (252).

By accumulating mutations during its natural history, the human HBV has evolved into at least eight recognised genotypes A–H (145, 146, 209-211, 217, 252, 270) and a putative genotype I (106, 125, 219, 294). Furthermore, a potential genotype J has recently been proposed (284). With the exception of E, G and H, genotypes are divided into subgenotypes with more or less distinct geographic distributions (6, 75, 107, 116, 117, 119, 120, 133, 142, 145-147, 150, 152, 202, 209, 220, 221, 223, 273, 274, 294). In addition, genetically close variants of HBV were also found in nonhuman primates, including chimpanzees (ChHBV), orangutans (OuHBV) and gibbons (GiHBV), each with its own distinct genotype (101, 179, 209, 249, 252). While cross-species transmission of HBV from humans to baboons (139) and gibbons (19, 258) was shown experimentally, there is no evidence of systematic natural transmission of HBV from non-human primates to humans (249).

1.1 Hepatitis B virus structure

Three different structures associated with HBV are found in the blood of infected individuals (Figure 1, Figure 2). Only the larger structure, the Dane particle, with a size of 42 nm is a functional viral particle (131, 244, 245). The small spheres (22 nm) and filaments (variable length), although usually found in 10,000 to 1,000,000-fold excess over the Dane particles, do not contain genomic material and are not infectious (306). The outer envelope of the virus, as well as the spheres and filaments consist of hepatitis B surface protein (HBs), also known as hepatitis B surface antigen (HBsAg). The HBs protein consists of three different forms: the small HBs protein (SHBs), the middle HBs protein (MHBs) and the large HBs protein (LHBs) (see chapter 1.2, p.3) (108). Depending on the size, the HBs protein consists of different domains. While the

SHBs solely contains the S domain, the MHBs and LHBs include the preS2 or preS1 and preS2 domains, respectively (Figure 2) (98). In addition, the virus particle harbours an inner capsid that is formed by 240 HBV core (HBc) protein units (66). The capsid surrounds the 3.2kb partially double stranded DNA genome that is covalently linked to the viral polymerase. Furthermore, each viral particle contains a protein kinase and the heat shock proteins hsp70 and hsp90 (Figure 2) (128).

1.2 Hepatitis B virus genome and gene products

The HBV genome is a partially double stranded, relaxed-circular DNA molecule of approximately 3.2kb (Figure 3), which replicates its genome via an RNA intermediate (108). Its circularity is maintained by the 5' cohesive ends of both the negative and positive strand and two short (10 or 11 nucleotides) direct repeats (DRs) in the viral DNA, a region that is suspected to also be important for integration into the hepatic genome (72, 306). The 5' end of the minus-strand is located in the DR1, while the plus strand starts at the DR2. The repeats are involved in priming the synthesis of the respective DNA strands (108, 306).

The minus strand is unit length, with the primase domain of the polymerase covalently linked to its 5' end, while the plus strand is less than unit length, with a capped oligoribonucleotide molecule at its 5' end (24, 155, 259).

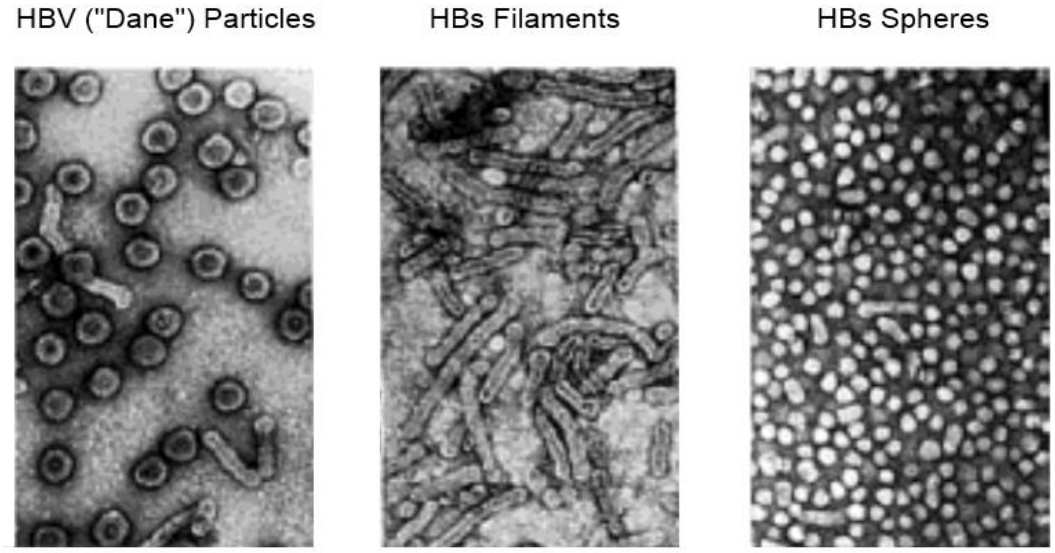

Figure 1: Electron-microscopy images of HBV particles.
Dane particles, filaments and small spheres can be observed in the serum of infected individuals (Figure adapted from (128)).

The complex structure of the HBV genome is furthermore increased by its small size and compact organisation, with four overlapping open reading frames (ORFs), polymerase (P), surface (S), X and core (C), coding for seven proteins. The starting point of the genome is defined as the *EcoRI* restriction site that is located on the S and polymerase ORF (Figure 3).

1.2.1 The polymerase (P) ORF

The polymerase (P) ORF is the longest HBV ORF and overlaps, at least partially, with all other HBV ORFs (Figure 3). The polymerase has at least four domains, with the N-terminal primase domain linked to the 5' end of the negative DNA strand (24, 25, 204). It is separated by a spacer domain from the RNA-dependent DNA polymerase, which has similarities to the reverse transcriptase of retroviruses (25, 26, 204, 291). The C-terminal domain encodes an RnaseH activity that cleaves the RNA of the DNA-RNA hybrid in the assembled viral capsids (100, 204).

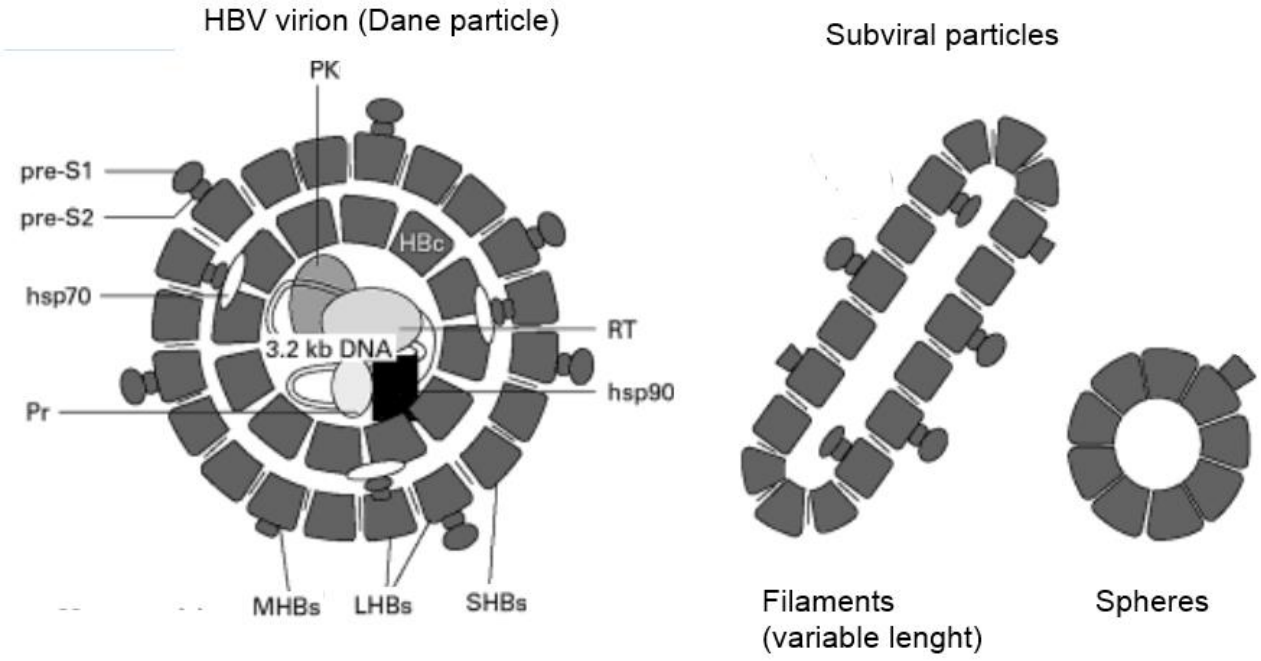

Figure 2: Schematic representation of HBV particles.
Empty spheres and filaments are secreted in the blood stream of infected donors in addition to the virus particle. The HBV virion consists of a HBs protein envelope and a HBc protein capsid. The encapsidated 3.2 kb DNA genome is covalently linked to the primase domain (Pr) of the DNA polymerase. Protein kinase (PK) and heat shock protein hsp90 are also encapsidated, while hsp70 is associated with internal preS domains of the LHBs. The figure is adapted from (128).

1.2.2 The surface (S) ORF

The surface (S) ORF codes for three different surface proteins, the small SHBs, the middle MHBs and the large LHBs, depending on the usage of one of the three in-frame translational start codons (Figure 3). For all three proteins, translation is terminated at a common stop codon. While the SHBs solely contains the S domain, the MHBs and LHBs include the preS2 or preS1 and preS2 domains, respectively (Figure 2, Figure 3). Consequently, the proteins share a common C-terminal region of 226 amino acids (AA). The MHBs contains an additional 55 AA at the N-terminus, while the LHBs is further extended N-terminally by 108 AA.

The common S-domain is generally referred to as HBV surface antigen (HBsAg) and is the main target of the host immune response during acute HBV infection. Anti-HBs-antibodies against important epitopes of the HBsAg, especially against the major antigenic determinant 'a' (amino acids 124–147), are critical for protection (330).

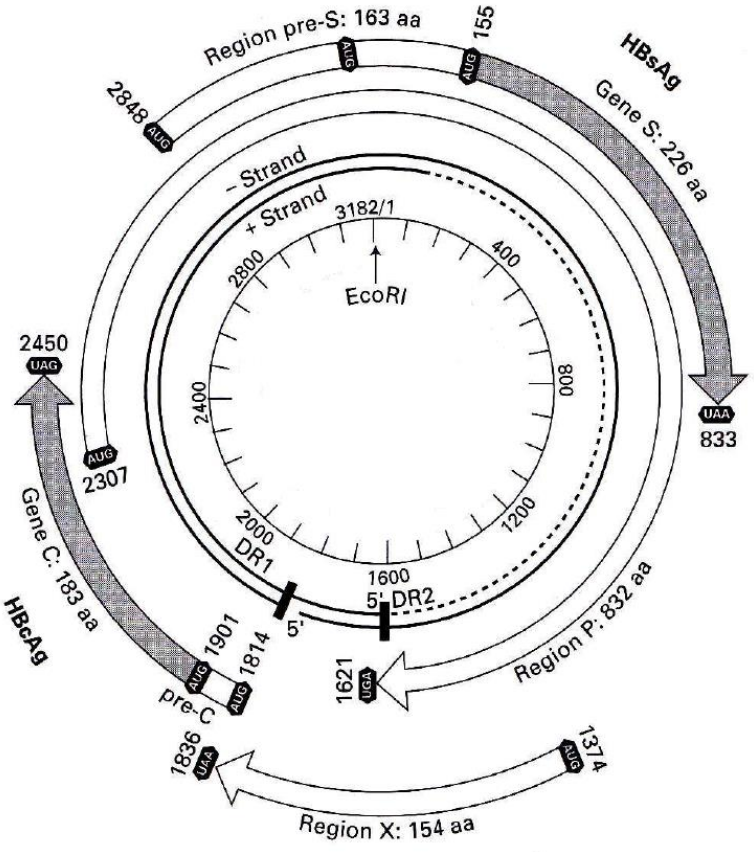

Figure 3: Organisation of the HBV genome.
The HBV genome is a partially double stranded DNA molecule of 3.2 kb size with four overlapping open reading frames, encoding seven transcripts (Figure adapted from (46)).

1.2.3 The core (C) ORF

The core ORF codes for two proteins, using two different translational start codons (69). The HBc protein or HBV core antigen (HBcAg) is 183 to 185 AA in length, depending on the viral genotype (123). The HBc protein assembles into the viral capsid of the infectious virion (69).

The HBe protein (HBeAg) is encoded by the same ORF, albeit from a different RNA species (see1.3 Life cycle of hepatitis B virus, p.7), using an alternative upstream start codon, thus including the N-terminal preC domain into the protein (69). The preC domain is of 29 AA length of which the first 19 AA encode a hydrophobic alpha helix that allows for translocation of the HBe protein into the lumen of the endoplasmic reticulum (ER), where the 19 AA are cleaved off by a signal peptidase (45, 269). The remaining ten residues are sufficient to prevent assembly of this protein. Following the transport through the ER and the Golgi apparatus, parts of the HBe protein undergo additional cleavage, and a heterogeneous population of HBeAg proteins, with a size of 15 to 18 kDa, is secreted (269).

1.2.4 The X ORF

The X ORF is only found in the genome of the *Orthohepadnaviridae* and codes for the X protein. It has a length of 154 AA (305). The function of the HBV X protein is not well understood, but it seems to be essential for the infection and replication of HBV. Indeed, studies in the woodchuck model (56, 328) or in mouse models (137, 138, 313) indicate that HBx is needed for productive infection. In addition, cell culture models were used to understand the function of HBx. While HBx deficiency had only a little effect when using Huh7 hepatoma cells, a key position of HBx in the initiation and maintenance of HBV replication was observed in HepG2 and HepaRG cells, as well as in primary human hepatocytes (34, 137, 138, 161, 175, 305). Furthermore, HBx was found to be a transactivator of multiple cellular and viral promoters (305), while some studies suggested that HBx activates viral transcription (40, 161, 326), or enhances encapsidation (190). However, mutational studies also suggest, that the X protein is not of relevance for the HBV life cycle (34). In addition, multiple studies observed tumorigenicity in model systems and truncated forms of HBx were observed in chronically infected individuals and suspected to play a role in HBV-related liver oncogenesis (236, 253, 271). Thus, despite the large number of studies performed, the function of HBx in HBV replication is not yet completely clear.

1.3 Life cycle of hepatitis B virus

During the last two decades the use of various hepatoma cell lines has profoundly increased the knowledge of the life cycle of HBV. However, especially the study of viral attachment has been hampered, as available *in vitro* systems such as HepG2 or HuH7 cell lines were dependent on transfection with the virus (98). Consequently, primary human hepatocytes (PHH) and those of chimpanzees, which were both limited in availability, were used for HBV infectivity studies (98). Recently, additional model systems have become available for the study of viral entry: primary hepatocytes from tree shrews (*Tupaia belangeri*) are used to substitute PHH and a new human hepatoma cell line (HepaRG) can be infected with hepatitis B virus (97, 102). A large number of studies on the life cycle of HBV were furthermore performed using two HBV-related animal viruses, the duck hepatitis B virus (DHBV), infecting Peking ducks and the woodchuck hepatitis virus (WHV), infecting the woodchuck *Marmota monax* (94, 183, 191, 255, 276).

Despite the limitations of the model systems, a large number of studies have been performed to understand the life cycle of HBV. Briefly, it can be divided into a number of steps (Figure 4) (128): (1) attachment of the infectious virion to the host cell membrane, (2) virus entry into the cell, (3) release of the viral core, (4) release of the viral genome into the nucleus, (5) transcription and translation of HBV genes, (6) genome replication, (7) virion assembly, (8) release of the virus. A schematic view of the HBV life cycle is presented in Figure 4.

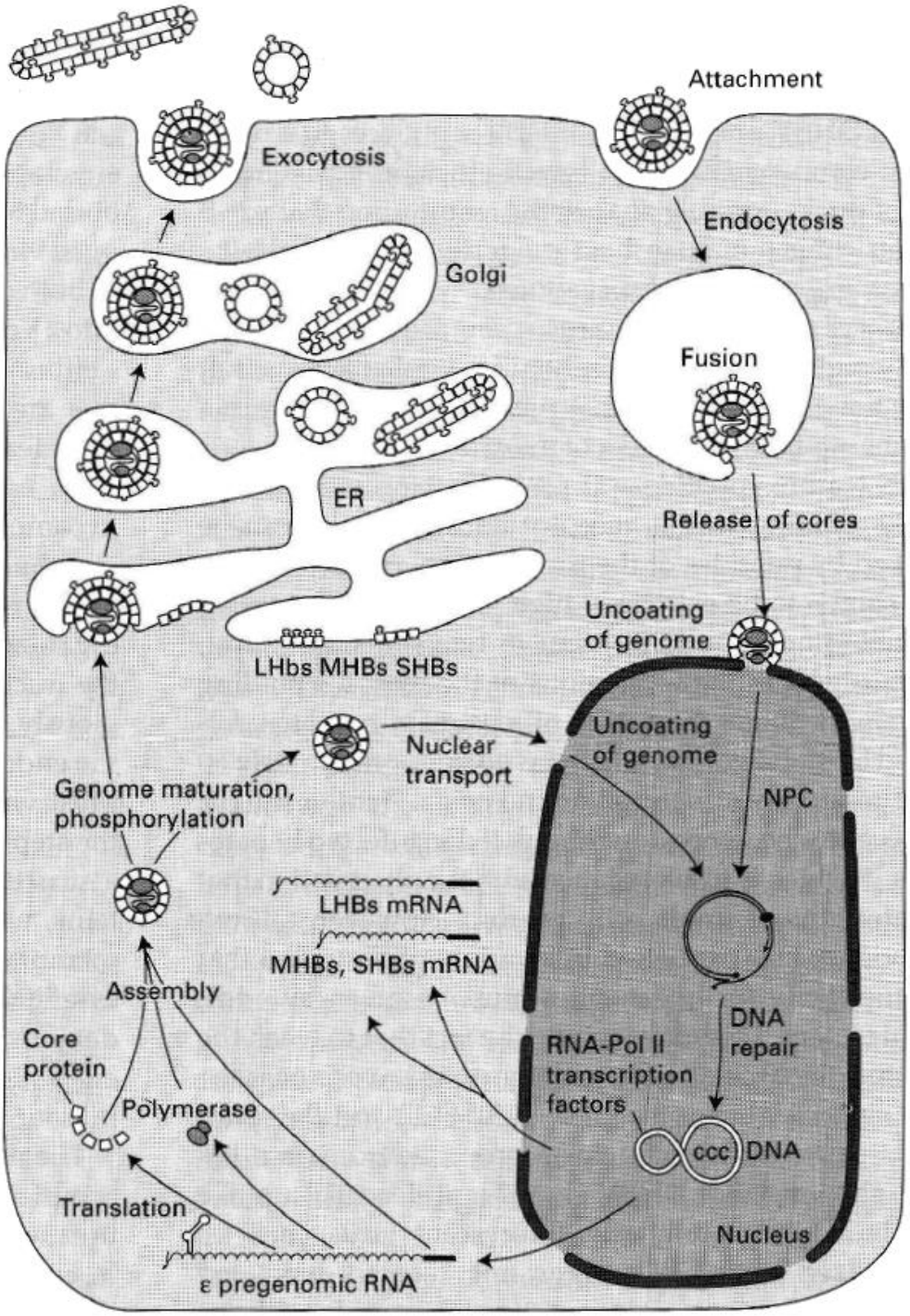

Figure 4: Schematic view of the HBV life cycle.
After attachment, HBV enters the hepatocyte, possibly by endocytosis. As soon as the nucleocapsid is released, it is transported to the nuclear pore, where the relaxed-circular (rc) DNA genome is converted into covalently closed circular (ccc) DNA. cccDNA is transcribed into three viral mRNAs of 2.4, 2.1 and 0.7 kb size and pregenomic (pg) RNA of about 3.5 kb size. The synthesis of the non-essential HBe and HBx are omitted in this figure. The core and polymerase transcripts are assembled in the cytosol, forming core particles with the pgRNA. The transcripts of the surface proteins are translated through the rough ER. The surface proteins are inserted in the ER membrane and bud into the ER lumen, with a subset of these enveloping the core particles. The HBV and HBs particles are then secreted constitutively. The figure is from (128) and is still speculative in many details.

The first step of HBV infection is the viral attachment to the cell surface, which appears to involve a multistep process (257). A cell type unspecific primary attachment to heparan sulfate proteoglycan is followed by a highly specific attachment step, involving the HBV surface proteins (257). Numerous studies reported a variety of interaction partners for all three HBV surface proteins, although recent studies point towards the preS1 domain as a key factor in mediating hepatocyte binding (98, 188, 207, 254, 257). In addition, both the differention status of the hepatocytes and the cell polarisation seem to have a large impact on the infection process (98, 256). Upon binding to the membrane, entry occurs probably via endocytosis, but the mechanism remains unclear (98). After the release into the cytoplasm, the nucleocapsid is transported to the nuclear pore and the partially double stranded relaxed-circular DNA genome (rcDNA) is released into the nucleoplasm and converted into a stable, covalently closed circular form (cccDNA) (128, 204). The episomal cccDNA then serves as the central transcriptional template to produce various subgenomic RNAs encoding the HBV proteins (Figure 4, Figure 5).

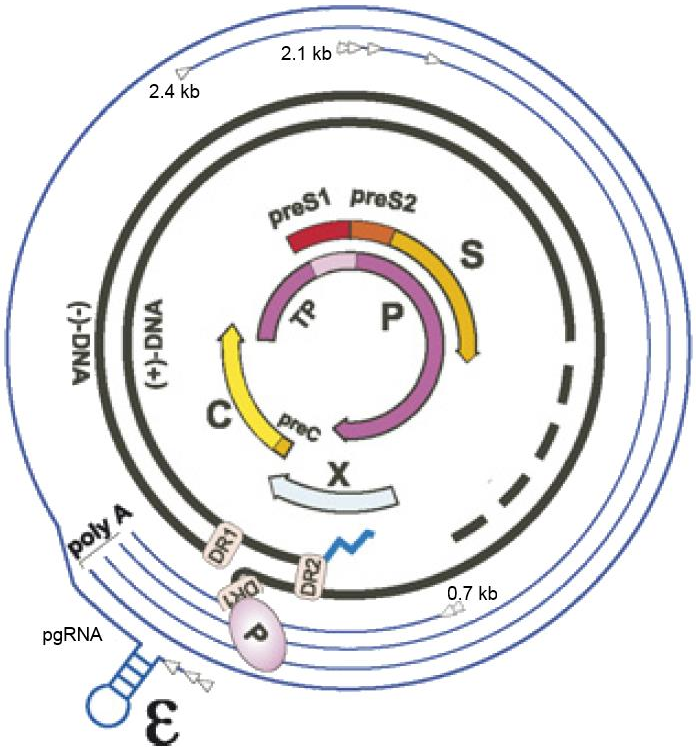

Figure 5: HBV genome and associated viral RNAs.
The inner circle represents the HBV DNA. The outermost lines depict viral RNAs present in infected cells. The three mRNAs are of 0.7 kb (encoding the X protein), 2.1 kb (middle and small HBs) and 2.4 kb (large HBs) size. The greater than genome length pgRNA encodes the HBc, HBe and polymerase and serves as template for HBV replication. The packaging signal (ε) is depicted as a hairpin structure. Figure adapted from (204).

The three viral mRNAs function as template for the viral proteins, using the cellular RNA polymerase II, with the 2.4 kb and 2.1 kb encoding the large or middle and small HBs protein, respectively, while the 0.7 kb mRNA encodes the X protein (128, 204, 243).

Two greater than genome length pregenomic (pg) RNAs of about 3.5kb encode the HBc, HBe and polymerase and serve as template for HBV replication (204). From the longer pre-core mRNA an extended non-assembling version of the core protein is translated that has an N-terminal extension with a hydrophobic alpha helix, allowing for the processing and secretion of the HBeAg via the ER (see 1.2.3 The core (C) ORF, p. 6). Another, slightly shorter, pgRNA serves as mRNA for the core protein and the polymerase (204). The newly translated polymerase then binds to the packaging signal ε at the 5'end of the pgRNA, preferentially to the molecule from which it was translated, and initiates the assembly with the core proteins to form a new nucleocapsid, in which the RNA is reverse transcribed into rcDNA (23, 204). The progeny capsids can then re-deliver their genomes into the nucleus to establish a pool of 10 to 100 copies of cccDNA or they bud into the ER to acquire their HBs envelope. Enveloped virions are then secreted through the constitutive pathway (44, 128, 154, 204, 295, 304, 325).

2 Clinical features of hepatitis B virus infection

HBV infection patterns vary in different subpopulations worldwide. This is largely influenced by the age at which infection occurs. HBV infections tend to be mild or asymptomatic in children, but 90% of infants and 30–50% of toddlers become chronic carriers (82, 110, 306). In contrast, acute hepatitis B infection is cleared from the bloodstream in more than 90% of adult cases (306). The clinical manifestations vary in both the acute and the chronic disease. The acute phase can range from a subclinical to an icteric hepatitis and even to fulminant hepatitis in some cases. Also chronic carrier state ranges from asymptomatic to cirrhosis and hepatocellular carcinoma (HCC).

2.1 Acute infections

After HBV transmission, the incubation period ranges from 1 to 4 months, but may be shorter if exposed to higher viral loads (22, 36). Symptoms of HBV infection are typically arthralgia and arthritis, i.e. pains and inflammation of joints in hands, arms

and shoulders, as well as skin rash and fever. These symptoms usually cease with hepatitis onset (36). The majority of infections continue without additional symptoms, while only subsets of them are accompanied by jaundice or unspecific symptoms.

The diagnosis of acute HBV is usually dependent on the presence of HBsAg and anti-HBc IgM in the serum of the patient. The viral HBsAg is the hallmark of HBV infection and the first serological marker to appear in the serum, followed by HBeAg, HBV DNA, and anti-HBc (Figure 6). IgM anti-HBc antibodies appear early in infection and are detectable for about 6 months, while IgG anti-HBc antibodies appear later and can persist for many years. Furthermore, levels of serum liver enzymes ALT (alanine amino transferase) and AST (aspartate amino transferase) increase, varying from a moderate increase of 3- to 10-fold to more than 100-fold (306). In acute, self-limiting hepatitis seroconversion of HBeAg and HBsAg to anti-HBe and anti-HBs and loss of HBV DNA from the serum occur within about six months following exposure.

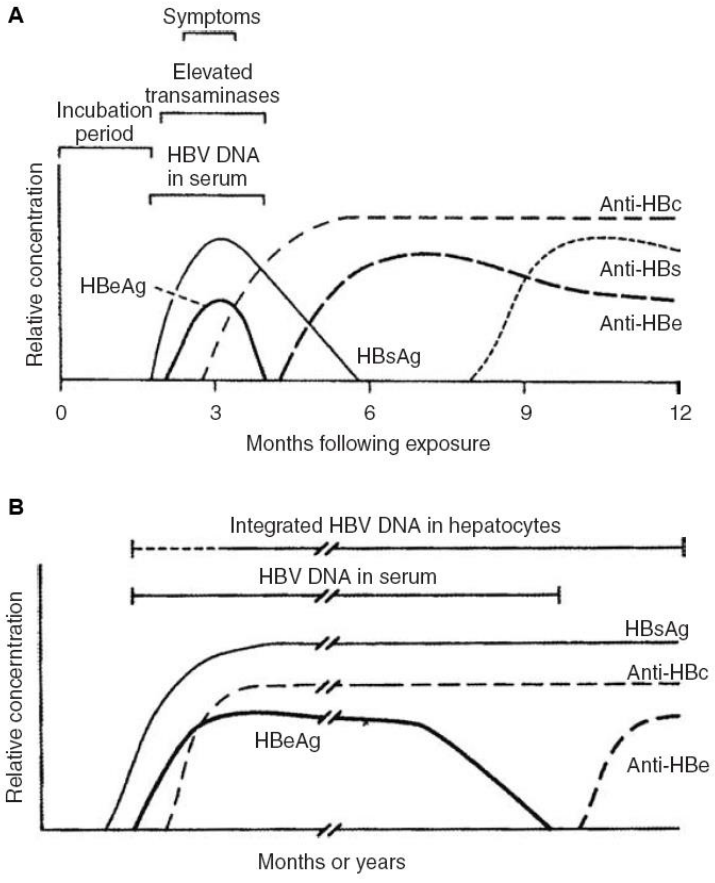

Figure 6: Serological profile of acute (A) and chronic (B) hepatitis B.
In acute, resolving infections HBV clearance is marked by the disappearance of HBV DNA and all HBV antigens and the presence of corresponding antibodies within 6 months after infection. Chronic infection is characterized by persistence of infection over a longer period of time. Figure adapted from (108).

Until recently, it was assumed that after recovery from acute hepatitis B, the virus would be cleared from the body. However, recent evidence suggests, that HBV persists as cccDNA in the hepatocytes of infected individuals, indicating that immunosuppression could lead to reactivation of the virus (36).

2.2 Chronic infections

In a large number of infants and toddlers, as well as 5–10% of adults, replicative HBV infection persists for more than 6 month, thus progressing to a chronic carrier state (90, 187, 306). Chronic carriers are characterized by the continuous presence of HBsAg, HBeAg and anti-HBc, which can persist throughout infection (Figure 6). HBeAg is an important marker of the replicative activity of the virus and seroconversion from HBeAg to anti-HBe indicates disease remission. Although the course of chronic infection can be highly variable between individuals, it can generally be divided into three phases: "immune-tolerant", "immune-clearance" and "inactive carrier" (166).

The immune-tolerant phase is characterized by HBeAg seropositivity, high levels of HBV DNA (>10^7-10^8 copies/ml) and normal ALT levels. Histological changes are rare during this stage, probably as a consequence of the host immune tolerance to HBV. This phase is more frequent and prolonged in patients infected perinatally or in the first years of life (81). Furthermore, transplacental transfer of maternal HBeAg is suspected to induce helper T-cell unresponsiveness in neonates (60). Thus, especially in high endemic countries with a late seroconversion of HBeAg to its antibody anti-HBe, patients in the immune-tolerant phase are usually young (166).

Transition from immune-tolerance to immune-clearance in patients infected perinatally or during early childhood varies in different geographical regions, where different genotypes prevail. While HBeAg seroconversion in Africa where genotypes E and A prevail, occurs before the age of 15 or 16, seroconversion in Asian carriers of genotype B and C occurs much later in life, approximately at the age of 30 and 40, respectively (60, 130, 319). The immune-clearance phase may last from several weeks to years and is characterized by decreased levels of HBV DNA (<10^7-10^8 copies/ml), abnormal and fluctuating ALT levels and increased histological changes (81).

HBeAg seroconversion is followed by the inactive HBV carrier phase with low or undetectable levels of HBV DNA (<10^4-10^5 copies/ml) and normal ALT. While patients are usually asymptomatic during this phase, periodic reactivation with fluctuating

levels of HBV DNA and ALT occurs in some patients. Also HBsAg seroclearance was found to increase with age, approximately from the age of 30 and substantially increasing after age 50 (61). Generally, low-level HBV replication persists in the liver with undetectable HBV DNA in the serum, while anti-HBc antibodies, with or without anti-HBs antibodies, remain detectable.

2.3 Occult infections

After the loss of HBsAg (i.e. a lack of detection with currently available assays), low-level HBV replication may still occur in a subset of chronically infected individuals, with detectable HBV DNA in the liver or blood, usually <1000 copies/ml or <200 IU/ml (242). The patho-physiological basis of occult hepatitis B infection (OBI) is still poorly understood but is probably due to an incomplete immune control of the infection. Most individuals with OBI have antibodies against HBV core antigen (anti-HBc), possibly reflecting an ongoing chronic infection with HBsAg below detection limits. This is particularly common in endemic countries where individuals are infected perinatally or during early childhood. OBI with antibodies against HBV surface antigen (anti-HBs), with or without anti-HBc antibodies, sometimes occurs during the recovery phase of the infection. Rare seronegative cases of OBI occur in individuals with no markers of HBV infection other than HBV DNA. Individuals with chronic hepatitis who have mutated HBsAg that is not detected by diagnostic assays also contribute to OBI (242).

3 Epidemiology

Worldwide, more than 2 billion people are or have been infected with hepatitis B virus at some time during their lives (308), accounting for approximately 600,000 deaths per year, most of which occur in the developing world (147, 197, 308). About 360 million people are estimated to be chronically infected and are at risk of liver cirrhosis and death due to liver failure (308). Furthermore, HBV is the major cause of hepatocellular carcinoma (HCC) worldwide (156), with the highest prevalences of HCC observed in regions with high HBV prevalence (30, 39, 43, 299, 317).

3.1 Global distribution of HBV infection

The prevalence of HBsAg, the primary marker of HBV infection, varies widely in different populations worldwide (Figure 7). Low prevalences (<2%) are found in North, West and Central Europe, North America and Australia (Figure 7) with prevalences

even below 0.5% in most regions (108). Intermediate prevalences of 2-7% are reported from the Mediterranean, Eastern Europe, Russia, Central and South America, and Southwest Asia, while high prevalences of ≥8% were largely observed in sub-Saharan Africa, Southeast Asia and parts of South America (Figure 7). In high risk populations the prevalence can be much higher, e.g. 55% in medical students in Mali (197) or more than 90% in HIV carriers in Cameroon (197). Also, the prevalence of anti-HBc antibodies is high and can reach over 85% in highly endemic regions (147).

3.2 HBV serotypes and genotypes

As the viral polymerase lacks proofreading capability, HBV shows a substantial genetic variability. Prior to the definition of genotypes, viral strains were classified by the four serotypes or HBsAg subtypes *adw*, *ayw*, *adr* and *ayr*. Classifications were based on the two mutually exclusive pairs *d*/*y* and *w*/*r*. The presence of lysine or arginine at AA 122 of the HBsAg indicates *d* or *y*, respectively, while lysine or arginine at AA 160 confer *w* or *r*. Furthermore, AA 113, 134, 144 and 145 sub-determine *d*/*y* specificity (52).

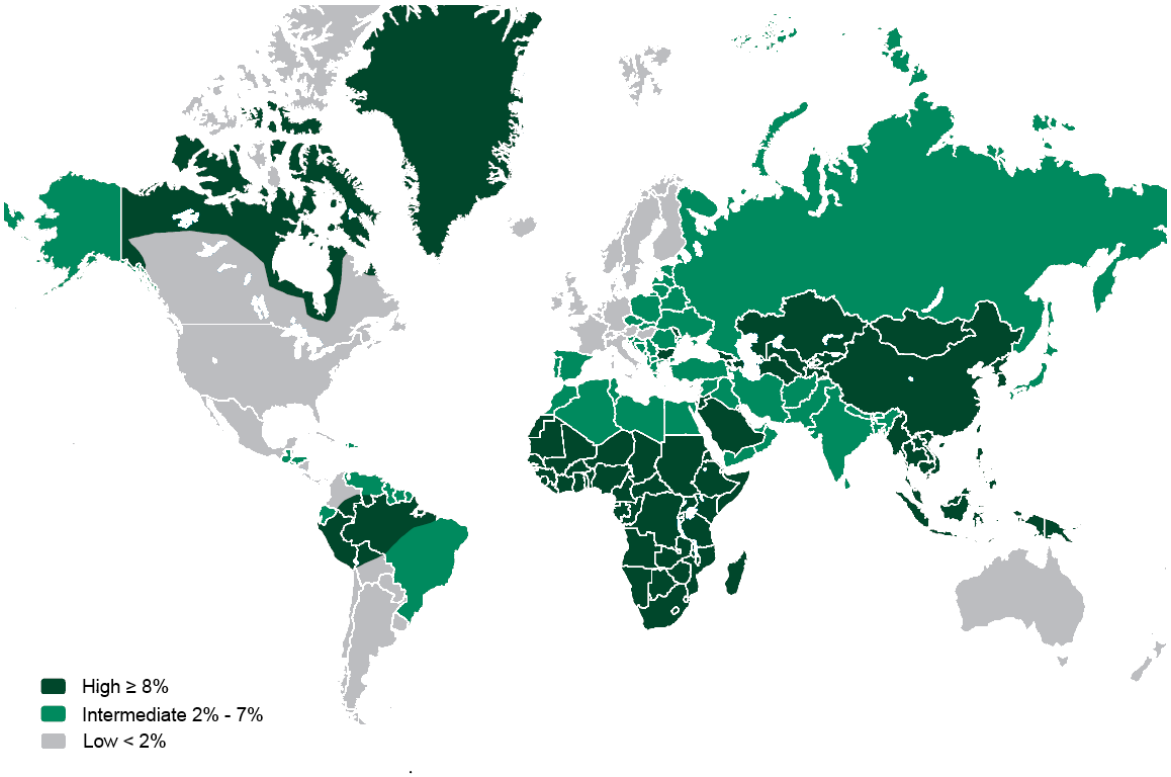

Figure 7: HBsAg prevalence worldwide.
Figure adapted from http://wwwnc.cdc.gov/travel/yellowbook/2012/chapter-3-infectious-diseases-related-to-travel/hepatitis-b.

In addition, restriction fragment length polymorphism (RFLP) was, and still is, used in settings where no access to sample sequencing is available. Restriction enzyme digestion of HBsAg PCR products and separation of digested fragments allow a determination of the serotypes and genotypes A-H (103, 321).

Today, amplification and sequencing has become the major tool for the analysis of HBV variants worldwide and to reconstruct the phylogenetic history of the virus. While genotypes were first defined based on a 8% nucleotide divergence of the complete genome or 4% on the S gene (217), the nucleotide divergence of the full-length genome was recently revised to 7.5% (145, 146, 209, 217). Additionally, subgenotypes are defined on a nucleotide divergence of 4% over the complete genome.

Based on the nucleotide divergence, HBV is classified into at least 8 recognised genotypes (146, 209). In addition, a putative genotype I was described, while a tenth genotype J has recently been proposed for a virus isolated from a Japanese patient (106, 220, 284, 294) (Figure 8). With the exception of genotypes E, G, H and J, genotypes are divided into subgenotypes with more or less distinct geographic distributions (Figure 8).

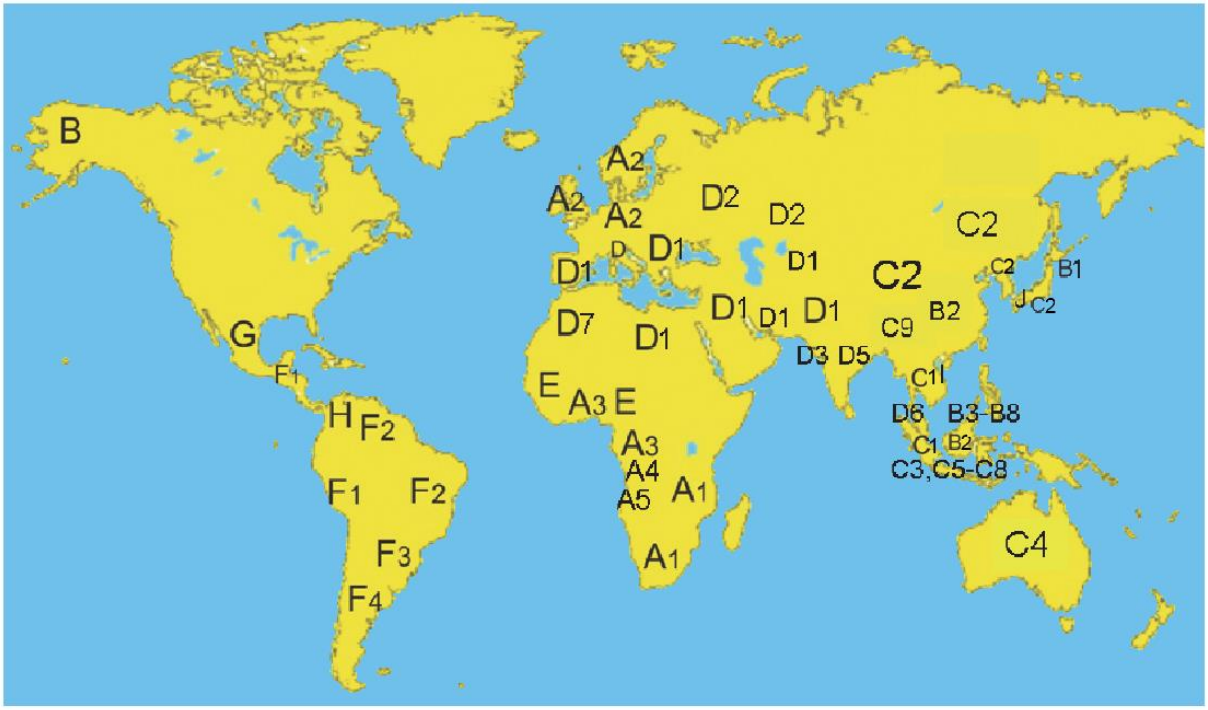

Figure 8: Worldwide distribution of HBV genotypes and subgenotypes.
Figure modified from (323).

For genotype A seven subgenotypes have been reported so far. A1 is prevalent in South and East Africa, Asia, India and Indonesia, but also in Brazil and Argentina (6, 21, 42, 107, 148, 150, 196, 199). A2 is mostly reported from Northwest Europe, but has also been detected in Africa, especially South Africa (142, 187, 209, 224, 252, 272). A3 was mainly found in West African countries, such as Cameroon and Gabon and The Gambia, but also in Rwanda (107, 116, 117, 152, 197). Of subgenotype A4 only a few strains were found in Mali and The Gambia (107, 221). A5 strains were first reported from Nigeria, a region that was a former slave trading post, but has subsequently been reported from Haiti and Cameroon (8, 116, 221). The recently proposed subgenotype A6 strains originate from African–Belgian patients originating from Congo and Rwanda (235), while subgenotype A7 was proposed for strains from Cameroon and Rwanda (117).

Genotypes B and C largely dominate in Asia and are divided into a number of subgenotypes. While B1 (formerly Bj) is mainly found in Japan and represents the non-recombined B subgenotype (274), subgenotypes B2 to B4 (formerly Ba) largely originate from mainland Asia and are recombinants with genotype C. B2 is mainly found in China and Southeast Asia, B3 mainly in Indonesia and B4 in Laos, Vietnam, Cambodia and China (51, 209, 220). The subgenotypes B5 and B6 have been isolated in the Philippines and from the Canadian Inuit populations, respectively, and subgenotypes B7, B8 and B9 have been proposed in Indonesia (199, 212, 287). Also genotype C strains group into several subgenotypes. According to the nomenclature from Huy et al. (119), subgenotype C1 represents predominantly strains from mainland Asia (Laos, Vietnam, Thailand and Myanmar), while C2 is prevalent in Japan, Hong Kong, China and Korea (55, 119, 145, 209). Subgenotype C3 consists of strains from Oceania, C4 of strains exclusively from Australia and subgenotype C5 strains have been reported from the Philippines, Indonesia, Vietnam and Laos (54, 106, 220, 250). Additionally, subgenotypes C6 to C16 have been proposed recently for strains from the Philippines and Indonesia (54, 198, 200, 201, 220).

Of the eight recognised genotype D subgenotypes, D1 is highly prevalent in the Mediterranean and the Middle East, and was also found in India and Pakistan (17, 21, 209, 252). D2 and D3 prevail in East Europe, including Russia and the Baltic States, while D3 is also present in India, Pakistan, South Africa and Rwanda (17, 117, 209, 281). D4 was observed in South Africa, Rwanda, Somalia and Oceania (117, 209). D5 strains were reported from East India and from Morocco (15, 21, 227), while

subgenotype D6 was observed in Indonesia (176, 198, 200, 201). In addition, two HBV/D subgenotypes were observed in Africa. While D7 was observed in Tunisia subgenotype D8, resulting from recombination between genotypes D and E, was reported from Niger (1, 189).

In contrast, genotype E has been found only in Africa, with some rare exceptions on other continents in persons with a link to Africa. Genotype E is found almost exclusively throughout the vast expanses of a crescent, spanning some 6000 km from Senegal in the West to the Central African Republic in the East and Namibia in the South (27, 48, 147, 177, 197, 278, 301). With about 70–140 million chronic carriers in this region, genotype E may well be the most important HBV genotype worldwide (197).

Genotype G has been observed in the USA and in France (270, 300). HBV/F subgenotype F1 was found in Alaska, Mexico, Central America, Peru and Argentina, while F2 to F4 were solely observed in South America (74, 187, 209). In addition, genotype H was found exclusively in Central and South America (11-13).

The putative genotype I, with its subgenotypes I1 and I2, was found in Laos, Vietnam, Northwest China and India (9, 106, 220, 248, 294, 316), while a tenth genotype J has recently been proposed in a virus isolated from a Japanese patient (284) (Figure 8).

3.3 Mixed infections and recombinations

In highly endemic countries, particularly in Asia where different HBV genotypes and subgenotypes co-circulate, mixed infections of different HBV variants (18, 114, 162, 168) and recombinations have been described (38, 41, 67, 195, 220, 265, 303, 314). For example, mixed infections have been reported from Thailand (124) and in China almost 50% of infections were found to be mixed (114).

Recombination events do not seem to be negligible artefacts of HBV co-infections, as recombinant strains have become the dominant variant in certain regions. Interestingly, most genotypes seem to be the result of recombination events (37, 265, 279). For example subgenotype Ba, a B/C recombinant, prevails in major parts of mainland Asia, and a recombinant between genotypes C and D has become the predominant variant in Tibet (67). Furthermore, genotype I seems to be a recombinant between genotypes G, C and A, while genotype E has been suggested to be a recombinant between genotype D and another unknown genotype (41, 220).

3.4 Transmission of HBV

HBV has been detected in a variety of body fluids. However, only serum, sperm and saliva have been demonstrated to be infectious, while transmission via breast milk, urine, tears and other body fluids remains controversial (4, 19, 33, 263).

WHO considers perinatal and early childhood transmission to be the most important routes of transmission for HBV worldwide [9]. The capacity to produce HBeAg and seroconversion to anti-HBe antibody is thought to play an important role in the transmission and outcome of perinatally acquired HBV infections (48, 140). The risk of perinatal infection increases if the mothers are HBeAg positive and 80% of their infants become chronic carriers (140). In Asia, where genotypes B and C prevail, HBeAg seroconversion occurs later in life (see 2.2 Chronic infections, p.12). Thus women of childbearing age are likely to be HBeAg positive and to transmit the virus to the offspring (60, 130, 260, 306, 319). In sub-Saharan Africa on the other hand, a region where genotypes E and A prevail, HBeAg seroconversion occurs before the age of 15 or 16 (48, 88, 260, 306). Thus horizontal infection in early infancy is estimated to be as low as 1–8% in sub-Saharan Africa (48, 140, 246). Nevertheless, early vertical childhood transmission is thought to be the most important infectious route in Africa with most children being infected by the age of 5 years (48).

Furthermore, horizontal transmission remains an important route of infection. Sexual contacts as well as percutaneous exposures e.g. due to contaminated medical equipment or contaminated blood donations, as well as injection drug use or tattooing are important routes of HBV transmission (87, 99, 127, 156, 208, 266, 308). However, transmission might also occur during household contacts (156). In addition, arthropod vectors, such as bedbugs and kissing bugs, have been implicated in the spread of HBV, but transmission has never been proven (31, 264).

3.5 Prevention of HBV infection

Infection with hepatitis B virus can be prevented by vaccination. The major response of individuals vaccinated against HBV is against the common ‘a’ determinant (see 1.2.2 The surface (S) ORF, p. 5) of the HBsAg, protecting against all known HBV genotypes (108).

First generation vaccines are plasma-derived vaccines consisting of purified and inactivated subviral HBsAg particles. They are derived from HBsAg positive donors, which are free of detectable HBV DNA (108, 306, 329). While these vaccines are still

produced in Asia they have been replaced by recombinant vaccines in many countries (306). Recombinant vaccines are produced by expressing the cloned HBV surface proteins (S, preS1, preS2) in yeast or mammalian cells (306, 329).

WHO recommends an inclusion of HBV vaccination into routine immunization services worldwide to prevent chronic HBV infections (306). In regions with high endemicity, administration of the first vaccine dose occurs immediately after birth, preventing transmission from HBeAg positive mothers in up to 70%. This is even increased to 90% by simultaneously administering the hepatitis B immunoglobulin (HBIG) that confers passive immunity for several months (108, 329).

The minimum protective level is generally considered to be 10 IU/l and booster doses are given to maintain a protective immunity. Nevertheless, 5 to 10% of healthy, immunocompetent individuals do not produce anti-HBs antibodies (nonresponders) or remain below the minimum threshold (hyporesponders) (108, 329). The nonresponders remain susceptible to infection with HBV (329).

3.6 Treatment of chronic HBV infection

The treatment of chronic HBV infection aims at the sustained depression of HBV replication in order to improve life quality and prevent the progression of disease. Nevertheless, HBV infection cannot be eradicated completely, as viral genomes persist in the hepatocytes as cccDNA (81).

Depending on the response to treatment, different endpoints of therapy are considered (81). While the ideal endpoint of therapy is a sustained loss of HBsAg, which is associated with a complete HBV remission and favourable long term outcome, this is not possible in all infected individuals. As HBeAg seroconversion is associated with improved prognosis, this is considered a satisfactory endpoint in HBeAg positive donors. However, in HBeAg positive individuals who do not achieve seroconversion and HBeAg negative patients, therapy aims at suppressing the HBV replication and depressing serum DNA to undetectable levels by sustained treatment (81).

Two different types of drugs can be used to reduce HBV DNA and viral replication: (pegylated) interferon alpha and nucleotide/nucleoside analogues (NUCs) (81). As (pegylated) interferon induces an antiviral state by inhibiting immunomodulation and cellular proliferation, no viral resistance is observed (108). However, side effects

are frequent and the drug has to be administered subcutaneously (76, 81). Furthermore, in patients with decompensated, HBV-related cirrhosis, autoimmune disease and those with uncontrolled severe depression or psychosis, interferon is contraindicated (81).

Multiple NUCs are available, such as entecavir, tenofovir, adefovir and telbivudine (81, 108). As these NUCs directly target steps of the HBV replication process, fewer side effects are reported than when using interferon, even after extended use (76, 81, 108, 165). However, the main disadvantage of NUCs is the development of drug resistances and cross-resistances, although in different rates for the individual drugs (76, 81). As this may result in treatment failure or HBV breakthrough during therapy, drugs have to be considered carefully for individual patients and patients should be monitored.

4 Concurrent infection with other viruses

As hepatitis B virus, hepatitis D virus (HDV), hepatitis C virus (HCV) and human immunodeficiency virus (HIV) share similar routes of transmission, concurrent infections occur. These were found to be clinically and virologically important, especially in areas with high prevalences, and complicate the course of disease, generally resulting in more severe and progressive liver disease than HBV infection alone. Infection with multiple viruses may occur as co-infection (i.e. simultaneous acute infection with different viruses) and super-infection (i.e. successive infection with different viruses). If the infection date of one or multiple viruses cannot be determined, the infection is referred to as concurrent (57).

4.1 Hepatitis Delta (D) virus

Hepatitis D virus (HDV) is a satellite virus that is dependent on HBV for its propagation. Worldwide, about 15 million HBsAg carriers are estimated to be also infected with HDV (285). The co-infection with these two viruses results in fulminant hepatitis more frequently than HBV infection alone, whereas super-infection of HBV with HDV is associated with chronicity of HDV in up to 80% of carriers (310).

HDV is a spherical particle of 36 nm size. It consists of an outer envelope containing the three HBV surface proteins SHBs, MHBs and LHBs (see 1.2.2 The surface (S) ORF, p. 5) and a nucleocapsid of hepatitis D antigen (HDAg) molecules, surrounding the approximately 1.7kb negative-stranded circular RNA genome (277).

Like HBV, HDV is transmitted through exposure to infected blood or body fluids as well as sexually (247, 310, 311). Perinatal transmission is however uncommon (310). HDV prevalences vary throughout the world, with regions of high endemicity in sub-Saharan Africa, East Europe, the Middle East, parts of Asia and certain South Pacific islands (Figure 9). The HDV prevalence is generally highest in regions with high HBV prevalence. However, in some regions with high HBsAg prevalence, e.g. in South East Asia, infection with HDV is uncommon (199, 206).

The HDV RNA genome is highly variable and strains are assigned to eight major clades, with divergence up to 18% within and up to 36% between clades (158). HDV clade 1 is the most prevalent clade worldwide and prevails in Africa, the Mediterranean basin, the Middle East as well as in Central and Northern Asia (73, 86, 118, 157). Clades 2 and 4 have so far been described in East and North-East Asia, with Clade 2 also present in the Yakutia region in Russia (121, 122, 312). Clade 3 has only been detected in South America (53). Clades 5 to 8 on the other hand have recently been described in West Africa and individuals that migrated to West Europe from this region (158, 178, 241). In this region HBV is highly endemic (6, 104, 147, 178, 197) and early childhood transmission is thought to be the most important route of infection (48). While this results in a high proportion of chronic HBV carriers, high rates of subsequent HDV super-infections add considerably to the burden of chronic liver disease (307).

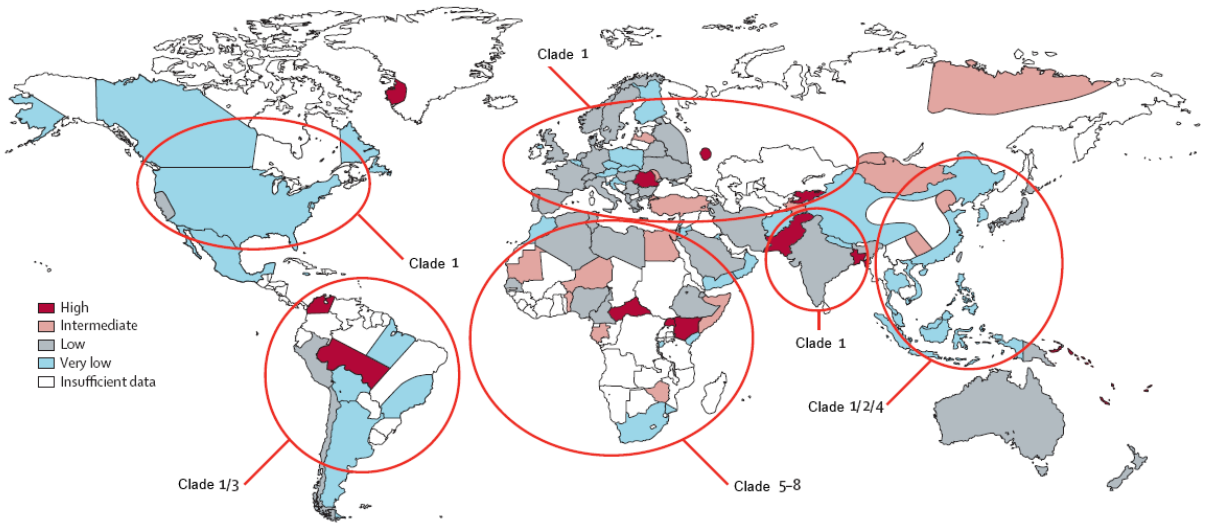

Figure 9: HDV prevalence worldwide and distribution of HDV clades.
Figure adapted from (118)

While HDV depends on the HBsAg for propagation, treatment regimens against HBV do not necessarily affect HDV replication. Today, (pegylated) interferon alpha is the only drug impairing HDV replication (81). Thus the prevention of infection is of utmost importance. While no specific HDV vaccine exists, vaccination of individuals who are not chronic HBV carriers against HBV provides protection against HDV infection (310).

4.2 Hepatitis C virus

In countries with high prevalences of perinatal or early childhood infection, the presence of both HCV and HBV is largely due to HCV super-infection of chronic HBV carriers. Worldwide, more than 10% of chronic HBsAg carriers are also infected with HCV (163). While super-infection with HCV in HBeAg and HBsAg positive carriers usually results in transient HCV infection, acute infection in HBeAg-negative, HBsAg-positive individuals advances to a persistent HCV infection. In addition, infection with both HBV and HCV seems to induce higher HBsAg seroconversion rates (63). Nevertheless, HCV super-infection is associated with earlier and more frequent progression to liver cirrhosis (164).

4.3 Human immunodeficiency virus

Human immunodeficiency virus (HIV) infection is prevalent in 10-15% of individuals chronically infected with HBV and is generally associated with increased levels of HBV DNA and decreased rates of spontaneous HBeAg clearance (62). Super-infection is furthermore associated with higher risk of liver related mortality than infection with HBV or HIV alone and chronic HBV carriers with HIV are at increased risk of liver cirrhosis (81, 289).

5 State of the art und objectives

Despite efficient vaccines, hepatitis B virus (HBV) continues to be a global public health problem. Especially in low income regions, such as sub-Saharan Africa and South-East Asia, HBV is highly prevalent, with prevalences reaching e.g. 55% in medical students in Mali (197) or more than 90% in HIV carriers in Cameroon (197).

In Asia, where HBV genotypes B and C prevail, HBeAg seroconversion occurs later in life and women of childbearing age are likely to be HBeAg positive and to transmit the virus to the offspring (60, 130, 260, 306, 319). Thus, perinatal transmission from the mother to the child is considered to be the most important route of infection in Asia, leading to 80% chronicity in infants (81, 140). In sub-Saharan Africa on the other hand, a region where genotypes E and A prevail, HBeAg seroconversion occurs largely before childbearing age, before the age of 15 or 16 (48, 88, 260, 306). Thus, vertical perinatal transmission and infection in early infancy is estimated to be as low as 1–8% in sub-Saharan Africa (48, 140, 246). Nevertheless, horizontal transmission during early childhood is thought to be the most important infectious route in Africa, with most children being infected by the age of 5 years (48).

In sub-Saharan Africa HBV is highly endemic. One of the two genotypes A and E dominates in most countries and only rare studies reported donors mixed infected with these genotypes. While genotype A is highly diverse with long evolutionary histories on the African continent, HBV/E, has a conspicuously low genetic diversity and seems to have a short evolutionary history. This, however, seems to be in contrast with its excessively high prevalence and its extensive spread throughout the vast West African genotype E crescent spanning from Senegal in the West to Namibia and the Democratic Republic of Congo in the South and East and including Ivory Coast, Ghana, Mali, Burkina Faso, Togo, Benin, Nigeria and Democratic Republic of the Congo, except for Cameroon where genotypes A and E co-circulate (27, 48, 147, 152, 177, 197, 221, 278, 301). With the exception of single sporadic cases with links to Africa (184, 267), HBV/E has not been found outside of Africa. It had been proposed that HBV/E evolved within the last 200 years (197). This would be compatible with the conspicuous absence of this genotype in Afro-Americans, despite the forced migration of slaves from West Africa to the New World during the transatlantic slave trade. Nevertheless, a recent study proposed a co-evolution of HBV with humans within the last 34,000 years and a time of evolution of HBV/E in Africa of 6000 years (229). However, if HBV/E has evolved from a single virus or a small

pool of viruses, it seems unlikely that the genetic diversity of about 1.7% would have taken such a long time to develop.

Furthermore, it is not known from where genotype E may originate and how it might have been introduced. While cross-species infections from humans to primates have been shown experimentally there is no evidence of natural HBV transmission. HBV/E has been found once in a chimpanzee, but the direction of transmission could not be established (280). Also arthropod vectors have been implicated in the spread of HBV, but transmission has not been proven (31, 264).

The wide spread and high prevalence of genotype E in Africa, despite its surprisingly low genetic diversity, represents an unsolved puzzle that is addressed in the first part of the present dissertation. First, molecular and phylogenetic analyses of genotype E and A strains from Africa and the Americas give a comprehensive picture of the spread and routes of transmission of HBV/E. The low genetic diversity of genotype E and its absence from the Americas is confirmed and implications on the time of evolution of genotype E are discussed (Chapter IV, Part 1). Additional analyses of the African HBV/E strains give valuable insights into the regional origin of this genotype in sub-Saharan Africa (Chapter IV, Part 2). Furthermore, phylogenetic analyses of HBV strains from Haiti, where more than 90% of the population descends from African slaves, and the inclusion of historical data further increase the understanding of the timeframe of evolution of genotype E and other West African HBV strains (Chapter IV, Part 3). In addition, extensive phylogeographic analyses confirm an only short time of evolution from a most recent common ancestor and refine the knowledge of the geographical origin of genotype E and the mutation rate of the virus, thus shedding light on the genotype E puzzle (Chapter IV, Part 4).

The burden of chronic liver disease is increased by the co-infection with other viruses, especially in highly endemic regions. In sub-Saharan Africa high prevalences of hepatitis D virus (HDV), a satellite virus of HBV, have been reported. Of the eight genetic HDV clades, four are found in Africa, with clades 5 to 8 having only recently been described in West Africa and individuals that migrated to West Europe from this region (158, 178, 241). As so far only rare and limited studies have been performed to assess the HDV prevalence in HBV chronic carriers (159, 213), we analyse more than 2000 samples from Burkina Faso, the Central African Republic, Chad and Nigeria. Our analyses shed light on the HDV burden in chronic HBV carriers in different cohorts, the

distribution of HDV variants and the characteristics of individual HDV clades (Chapter IV, Part 5).

Additional studies are performed in Lao PDR, where about 9% of blood donors are chronic carriers of HBV (126) and, like in large parts of Asia, the evolutionary 'older' genotypes B, C and I co-circulate(220). Genotypes B and C are often found at higher viral loads in the blood than other genotypes and are associated with seroconversion at a higher age than genotype E in Africa (48, 167). As a result, perinatal transmission of HBV is frequent in Asia, leading to a high proportion of chronicity. In highly endemic countries, where different HBV sub-/genotypes co-circulate, mixed infections (18, 114, 162, 168) and recombinations have been described (38, 41, 67, 195, 220, 265, 303, 314). Also, recombination events do not seem to be negligible artefacts of HBV co-infections, as recombinant strains have become the dominant variant in certain regions. For example, subgenotype Ba, a B/C recombinant, prevails in major parts of mainland Asia, and a recombinant between genotypes C and D has become the predominant variant in Tibet (67). Recently a recombinant variant that circulates in Laos and Vietnam has been proposed as a new HBV genotype I (106, 220, 248, 294). Nevertheless, only few systematic studies of HBV recombinants have been published from regions with intensive co-circulation of several HBV subgenotypes (220, 274, 303).

In this part of the thesis valuable results of clonal characterization, phylogenetic analyses and recombination analyses shed light on the impact of the co-circulation of HBV sub-/genotypes on the frequency of recombination events and the generation of new viral variants (Chapter IV, Part 6). In Lao PDR blood donations are, so far, screened for HBV by only testing for the HBV surface antigen. Thus, HBsAg-negative and DNA positive, i.e. occult infected, donors are currently not rejected. This chapter determines the serological profile of Lao blood donors, the prevalence of occult infections and discusses possible implications for the safety of blood transfusions in Lao PDR (Chapter IV, Part 7).

Chapter II: Materials

1 Chemicals and reagents

Compound	Company
Agarose	Lonza
Ampicillin	Sigma
Dimethylsulfoxide (DMSO)	Sigma
Dithiothreitol (DTT) 0.1 M	Invitrogen
Ethanol 96-100%	Merck
Ethidium bromide	Invitrogen
Ethylendiaminetetraacetic Acid (EDTA)	Biorad
Glycerol	Sigma
Hi-Di™ Formamide	Applied Biosystems
Kanamycin	Sigma
Luria Bertani Broth Base (LB)	Invitrogen
Magnesium Chloride ($MgCl_2$)	Invitrogen
Nucleotides (dNTPs)	Invitrogen
Oligonucleotides/Primers	Eurogentec
Orange G	Invitrogen
PCR Buffer without $MgCl_2$ (10x)	Invitrogen
PicoGreen® 10 000X	Molecular Probes
Potassium chloride (KCl)	Merck
Random hexamer	Invitrogen
Sodium Acetate	Merck
Sodium chloride (NaCl)	Merck
Sodium hydroxide (NaOH)	Merck
Sucrose	Sigma
SYBR®Green™ nucleic acid stain (10,000x)	Molecular Probes
SYBR®Safe™ DNA Gel Stain (10,000x)	Invitrogen
Tris(hydroxymethyl)aminomethane (Tris)	Sigma

2 Buffers and Solutions

Buffer/solution	Reagent	Volume/concentration
DNA loading dye (6x)	Orange G	25mg
	Sucrose (40%)	4g
	ddH_2O	fill up to 10ml
TAE-Buffer (50x)	Tris	2M
	Sodium Acetate	25mM
	EDTA	0.5M
	Adjust to pH 7.8	
S.O.C. Medium (Invitrogen)	Tryptone	2%
	Yeast Extract	0.50%
	NaCl	10mM
	KCl	2.5mM
	$MgCl_2$	10mM
	$MgSO_4$	10mM
	Glucose	20mM
Phosphate-Buffered-Saline (PBS)	NaCl	8g
	KCl	0.2g
	Na_2HPO_4	1.44g
	KH_2HPO_4	0.24g
	ddH_2O	fill up to 800ml
	Adjust pH to 7.4 with HCl	
	ddH_2O	fill up to 1l

3 Enzymes

Enzyme	Company
Phusion™ High Fidelity DNA polymerase	Finnzyme
Platinum® Taq DNA polymerase	Invitrogen
RNaseOUT™ (Recombinant Ribonuclease Inhibitor)	Invitrogen
SuperScript™ III Reverse Transcriptase	Invitrogen
TaqMan Universal PCR master mix	Applied Biosystems

4 DNA marker

1kb plus DNA ladder™	Life Technologies

5 Standard

2nd WHO International Standard for Hepatitis B Virus DNA, NIBSC code 97/750	NIBSC, Hertfordshire, UK

6 Bacterial strain

E. coli One Shot®TOP10	Invitrogen

7 Bacterial growth medium

Medium	Reagent	Volume/concentration
Luria Bertani Broth (LB)	Luria Bertani Broth Base	25 mg
	ddH_2O	fill up to 1l
	Autoclave at 121°C for 15 min	

To prepare growth plates, 32 mg LB agar were added per 1 l medium. Bacteria were plated onto Luria-Bertani-Broth agar plates, in the presence of kanamycin [30 mg/ml] and X-Gal.

8 Vectors

Vector	**Company**
pCR® 4-TOPO®	Invitrogen
pCR®-Blunt II-TOPO®	Invitrogen

9 Commercial Kits

Kit name	Company
AxSYM® HBsAg V2	ABBOTT Diagnostics
AxSYM® HBe 2.0	ABBOTT Diagnostics
AxSYM® AUSAB	ABBOTT Diagnostics
AxSYM® CORE	ABBOTT Diagnostics
AxSYM® Anti-HBe 2.0	ABBOTT Diagnostics
Big Dye® Terminator v3.1 Cycle Sequencing Kit	Applied Biosystems
ETI-AB-DELTAK-2	Diasorin
Jet Quick PCR Purification Spin Kit	Genomed
MagMAX™-96 Total Nucleic Acid Isolation Kit	Ambion
Murex anti-Delta	Diasorin (formerly distributed by ABBOTT Diagnostics)
Murex HBsAg Version 3®	Diasorin (formerly distributed by ABBOTT Diagnostics)
QIAamp Viral RNA Mini Kit	Qiagen
QIAamp® DNA Blood Mini Kit	Qiagen
QIAprep® Spin Mini Kit	Qiagen
QIAquick® Gel Extraction Kit	Qiagen
Smart Check HBsAg Test	GlobaleMed
TOPO TA Cloning® Kit	Invitrogen
Zero Blunt® TOPO® PCR cloning kit	Invitrogen

10 PCR primers

10.1 Hepatitis B virus

Table 2: Primers and conditions for the amplification and quantification of hepatitis B virus

Target region	Primer name	Primer orientation	Amplification round	5'-3' Sequence	Ref.	Primer [µM]	$MgCl_2$ [mM]	Annealing T[°C], t[s]	Elongation T[°C], t[s]
preS	fw2422	forward	first	agaactccctcgcctcgcagac	(221)	0.2	1.65	63, 20	72, 60
	preS-R	reverse		acaggcggkgtttttcttgttga	(221)				
	fw2451	forward	second	tcaatcgccgcgtcgcagaa	(221)	0.2	1.5	63, 20	72, 60
	preS-R	reverse		acaggcggkgtttttcttgttga	(221)				
S	P2f	forward	first	cctgctggtggctccagttc	(221)	0.2	1.65	63, 20	72, 60
	979	reverse		attggaaagtatgtcaaagaattgtgggtcttttg	(221)				
	P2f	forward	second	cctgctggtggctccagttc	(221)	0.2	1.5	60, 20	72, 60
	Mc2R	reverse		tggaagttggggatcattgcc	(221)				
X	455	forward	first	caaggtatgttgcccgtttg	(221)	0.2	1.5	62, 20	72, 60
	rv1800	reverse		agaccaatttatgcctacagcctccta	(221)				
	fw696	forward	second	tcagtggttcgtagggctttcc	(221)	0.2	1.5	63, 20	72, 60
	rv1800	reverse		agaccaatttatgcctacagcctccta	(221)				
C	fw1608	forward	first	gcatggagaccaccgtgaacg	(221)	0.2	2.0	60, 20	72, 60
	rv2661	reverse		tcatttacagtgagagggcccacaaattg	(221)				
	fw1644	forward	second	tgcccaaggtcttacataataggactcttg	(221)	0.2	1.8	60, 20	72, 60
	rv2661	reverse		tcatttacagtgagagggcccacaaattg	(221)				
Taqman	HBV-QTqm-F	forward	-	actcaccaacctcttgtcct	(171)	0.2	-	95, 15	60, 60
	HBV-QTqm-R	reverse		gacaaacgggcaacatacct	(171)	0.2			
	HBV-QTqm-Probe	forward		FAM-tatcgctggatgtgtctgcggcgt-TAMRA	(171)	0.1			

10.2 Hepatitis D virus

Table 3: Primers and conditions for the amplification and quantification of hepatitis D virus

Target region (nt)	Primer name	Primer orientation	5'-3' Sequence	Ref.	Phusion buffer †	Primer [µM]	$MgCl_2$ [mM]	DMSO	Annealing T[°C], t[s]	Elongation T[°C], t[s]
307-870	320ds	forward	ccagagramcccttcarcgaac	(122) *	GC	0.8	1.5	2%	66, 30	72, 35
307-870	rv900	reverse	gtccgacctgggcatccg							
307-870	n320ds	forward	ccagagramccctctccarcgaac	(122) *	GC	0.8	1.5	2%	66, 30	72, 35
307-870	rv900	reverse	gtccgacctgggcatccg							
715-1302	710s	forward	cgccggctgggcaacatt		GC	0.8	1.5	2%	66, 30	72, 35
715-1302	1302das	reverse	ggnttcaccgacraggagag							
868-483	fw900_2a	forward	gaccgcgrggaggtggagatg		HF	0.8	1.5	-	62, 30	72, 50
868-483	480as	reverse	ccgggataagcctcactc							
Taqman	Delta-F	forward	gcatggtcccagcctcc	(159)		0.3				
Taqman	Delta-R	reverse	tcttcgggtcggcatgg	(159)	-	0.3	-	-	95, 15	60, 60
Taqman	Delta-Probe	forward	FAM-atgcccaggtcggac-MGB	(159)		0.2				

* primer sequence modified; † GC and HF buffers are provided with the Phusion™ High Fidelity DNA polymerase

10.3 M13 cloning verification PCR

Table 4: Primers and conditions for M13 PCR

Primer name	Primer orientation	5'-3' Sequence	Primer [mM]	$MgCl_2$ [mM]	Annealing T[°C], t[s]	Elongation T[°C], t[s]
M13fw	forward	gtaaaacgacggccag	0.8	2.5	58, 30	72, 60
M13rv	reverse	caggaaacagctatgac				

11 Software

Software	Reference
Adobe Photoshop	Adobe Systems Incorporated
Adobe Illustrator	Adobe Systems Incorporated
BEAST v.1.6.2	(78)
BEAUTI v.1.6.2	(78)
BioEdit v7.0.9.0	http://www.mbio.ncsu.edu/BioEdit
BLAST	http://ncbi.nlm.nih.gov/BLAST
CFX Manager	BioRad
Clustal W	(290)
Data Collection Software v3.0	Applied Biosystems
FastPCR v3.7.8	(R. Kalender, University of Helsinki, Finland)
FigTree v1.3.1	(http://tree.bio.ed.ac.uk/software/figtree)
Google Earth v6.2.1	Google Inc.
jPHMM	(322)
LogCombiner v.1.6.2	(78)
MAFFT v6b	http://mafft.cbrc.jp/alignment/software (135, 136)
MEGA 4.0.2	http://www.megasoftware.net (282)
MEGA 5.0.5	http://www.megasoftware.net (283)
Microsoft Office Suite 2003	Microsoft Corporation
Microsoft Office Suite 2010	Microsoft Corporation
SigmaPlot	Systat Software, Inc.
SigmaStat	Systat Software, Inc.
Modeltest	(234)
Network v4.610	Fluxus Technology, Germany (20)
Network Publisher	Fluxus Technology, Germany (20)
Opticon Monitor™ v3.1	BioRad
RDPv.3.44	(182, 225, 268)
SeqScape® v2.5	Applied Biosystems
SimPlot v3.5.1	http://http://sray.med.som.jhmi.edu/SCRoftware/simplot (173)
SPREAD v1.0.3	(29)
SPSS 17.0 for Windows	SPSS Inc., Chicago, USA
Topali v2	(192)
Tracer v.1.6.2	(78)
TreeAnnotator v.1.6.2	(78)

12 Instruments

Instrument	Reference
Automated analyser	AxSYM® system, ABBOTT Diagnostics
Balance	SARTORIUS Precision Balance
Centrifuges	Pico 17, Heraeus® Biofuge Stratos, Heraeus® UNIVAPO 150H, UniEquip
Electroporation apparatus	Pulse Controller Plus, Capacity Extender Plus, Gene Pulser II Plus, Biorad
Electrophoresis power supply	E835, Consort
Fluorescence reader	GENios Plus, Tecan
Gel Tank and Casting Form	Biozyme
Gel Documentation System	InGenius, Syngene
Heating Block	Thermomixer Comfort, Eppendorf
Incubator	HERAcell® 150, Heraeus Binder Incubators, Binder
KingFisher Flex	ThermoScientific®, Thermo Fisher Scientific
NanoDrop ND-1000 Spectrophotometer	Isogen
PCR Machine	Mastercycler® Gradient, Eppendorf UnoCycler, VWR
PRO200 Rotor-Stator Homogenizer	PRO Scientific Inc.
Real Time PCR Machines	Opticon® 2 DNA Engine, Chromo4TM, CFX, MiniOpticon, Biorad
Safe Imager™ 2.0	Invitrogen
Sequencer	ABI PRISM® 3130xl Genetic Analyzer, Applied Biosystems
Shaker	Multitron 2, INFORS-HT
UV transilluminator	Safe Imager™ 2.0 Blue-Light, Transilluminator, Invitrogen
Vacuum source	Vacuum Pump, UNIequip
Vortex	Vortex-Genie® 2, Scientific Industries

Chapter III: Methods

1 Antigen and antibody detection

1.1 Enzyme-linked immunosorbent assay (ELISA)

1.1.1 HBsAg detection

The detection of the hepatitis B virus surface antigen (HBsAg) in the serum of voluntary blood donors and liver patients was performed using the Murex HBsAg kit version 3® and following the manufacturer's protocol. The assay is based on the enzyme-linked immunosorbant assay principle (ELISA). Briefly, 25 µl of sample diluent were added to the wells of a 96 well microtitre plate, followed by 75 µl of serum sample. On each plate, two negative controls and one positive control were included (provided with the kit). The plate was then covered by a lid and incubated at 37°C for 1 h. After incubation, 50 µl of conjugate were added to each well and the plate was incubated again at 37°C for 30 minutes. At the end of the incubation time, the plate was washed 5 times with wash fluid using an ELISA washer. 100 µl of substrate solution were added to each well. After incubation for 30 minutes at 37°C, 50 µl of stop solution (0.5 M sulphuric acid) were added to each well. The absorbance was measured using an ELISA reader at a wavelength of 450 nm with 690 nm as a reference wavelength.

1.1.2 HDV-Ab detection

The detection of hepatitis D virus antibodies (HDV-Ab) was performed using the Murex anti-Delta or ETI-AB-DELTAK-2 ELISA kit, following the manufacturers' protocols.

For the Murex anti-Delta kit, 25 µl of sample diluent were added to the wells of a 96 well microtitre plate followed by 75 µl of serum sample. On each plate, two negative controls and one positive control were included (provided with the kit). The plate was then covered by a lid and incubated at 37°C for 1 h. After incubation, 50 µl of conjugate were added to each well and the plate was incubated again at 37°C for 60 minutes. At the end of the incubation time, the plate was washed 5 times with wash fluid using an ELISA washer and 100 µl of substrate solution were added to each well. After incubation for 30 minutes at 37°C, 50 µl of stop solution (0.5 M sulphuric acid) were added to each well. The absorbance was measured using an ELISA reader at a wavelength of 450 nm with 690 nm as a reference wavelength.

For analyses with the ETI-AB-DELTAK-2 ELISA kit, 50 µl of sample were added to wells of a 96 well microtitre plate. On each plate, three negative controls and two positive

control were included (provided with the kit). In addition, a blank well without control/sample and enzyme tracer was included. 100 µl of diluted enzyme tracer were added to each well, except for the blank well. The plate was then covered by a lid and incubated at 37°C for 3 hours. At the end of the incubation time, the plate was washed 5 times with wash fluid using an ELISA washer and 100 µl of substrate were added to each well. After incubation at room temperature in the dark for 30 minutes, 100 µl of blocking reagent were added to each well. The absorbance was measured using an ELISA reader at a wavelength of 450 nm and using 630 nm as a reference wavelength.

1.2 Microparticle enzyme immunoassay (MEIA)

For our study on occult HBV infection in Lao PDR, all assays were performed on an automated AxSYM® system in partnering institutions. HBsAg, HBeAg and anti-HBs, anti-HBc, anti-HBe antibodies were detected by AxSYM® HBsAg V2, AxSYM® HBe 2.0, AxSYM® AUSAB, AxSYM® CORE and AxSYM® Anti-HBe 2.0 assays, respectively. Analyses were performed according to manufacturer's instructions.

1.3 Rapidtest

For some samples HBsAg detection was performed in partnering institutions, using the Smart Check HBsAg Test, according to manufacturer's instructions.

2 DNA extraction

DNA was extracted from serum samples using the QIAamp® DNA Blood Mini Kit, following the manufacturer's protocol. Briefly, 200 µl serum were added to 20 µl ProteinaseK and 200 µl lysis buffer (provided with the kit) in a 1.5 ml centrifugation tube. After vortexing, the resulting solution was incubated at 56°C for 10 min. After a brief centrifugation, 200 µl ethanol (100%) were added. After vortexing and centrifugation, the mixture was applied to a QIAamp® spin column (in a 2ml elution tube). The column was centrifuged at 6000g for 1 min. The flow through was discarded. The column was then washed in two centrifugation rounds with 500 µl of washbuffers AW1 and AW2, respectively. The DNA was eluted into a 1.5 ml tube by application of 60 µl elution buffer to the membrane, incubation for 1 min and centrifugation at 6000g for 1 min. Eluted DNA was stored at -20°C.

3 RNA extraction

RNA was extracted from serum samples using the QIAamp Viral RNA Mini Kit (Qiagen), following the manufacturer´s instructions. 140 µl of serum were added to 560 µl of AVL buffer (lysis buffer provided with the kit) containing carrier RNA, in a 1.5 ml cetrifugation tube. After vortexing, the resulting solution was incubated at room temperature for 10 min. To this 560 µl ethanol (100%) were added. After vortexing and brief centrifugation the mixture was applied to a QIAamp® mini spin column (in a 2ml elution tube). The column was centrifuged at 6000g for 1 min. The flow through was discarded. The column was then washed in two centrifugation rounds with 500 µl of washbuffers AW1 and AW2, respectively. RNA was eluted into a 1.5 ml tube by application of 60 µl elution buffer to the membrane, incubation for 1 min and centrifugation at 6000g for 1 min. Eluted RNA was stored at -80°C.

4 Polymerase chain reaction

The Polymerase chain reaction (PCR) is generally used to amplify double-stranded DNA in order to obtain a high amount of DNA for further experiments or to quantify a specific DNA fragment. Specific oligonucleotides (primers) are used to amplify a specific DNA region. The DNA amplification is facilitated by a DNA dependent DNA polymerase during multiple PCR cycles. Each cycle typically consists of a denaturation, annealing and elongation step.

In order to increase the sensitivity and specificity of the analysis, a semi-nested PCR setup can be used. In this approach, a product obtained by PCR is subjected to another round of amplification, using one first round primer and another primer that would bind towards the centre of the first-round product. This approach increases the sensitivity of the PCR and reduces the risk of unspecific amplification in the second-round PCR.

Besides template DNA, primers and polymerase, the mixture also contains dNTPs, $MgCl_2$ as well an enzyme-specific buffer. Furthermore, SYBR® Green, a fluorescent molecule that binds to double-stranded DNA, can be added to the mix for quantification in a real-time PCR machine.

In addition the TaqMan® assay, also known as 5' nuclease assay, is used for quantitative analyses. In addition to the primers, this assay uses an additional oligonucleotide (probe). The probe is coupled to a fluorescent molecule and binds

to the single DNA strand. The fluorescent molecule is released from the probe during elongation by the DNA polymerase and thus emits a fluorescent signal.

4.1 Amplification of the hepatitis B virus genome

The HBV genome was amplified in four overlapping PCR fragments, of approximately 1000 nt of size each, called preS, S, X and C fragments. A semi-nested PCR was used, except for the analyses of mixed infections and recombinations in Lao PDR. 5 µl extracted DNA or 1 µl first-round product were amplified in the first- and second-round PCR, respectively. Primers and variable conditions are listed in Table 2. Constant PCR parameters are listed in Table 5.

Table 5: Parameters for PCR using the Platinum® Taq DNA polymerase

Reagent	End concentration or volume
PCR Buffer	1x
$MgCl_2$	See Table 2 and Table 4
dNTP	200 nM
Forward primer	0.2 µM
Reverse primer	0.2 µM
SYBR® Green (if appropriate)	1x
Platinum® Taq DNA polymerase	1 unit
Template	5 µl (1st round) or 1 µl (2nd round)
RNase/DNase free H_2O	Add up to 25 µl total reaction volume

4.2 Reverse transcription of extracted RNA

Reverse transcription was carried out to reverse-transcribe RNA into DNA. Briefly, 5 µl of extracted RNA were denatured with 45 ng random hexamers and 10 nmol dNTPs for 5 min at 72°C. The reaction was then performed for 80 min at 50°C using 200U SuperScript III reverse transcriptase and 40U RNAseOUT recombinant Ribonuclease Inhibitor.

4.3 Amplification of hepatitis D virus genome

The hepatitis D virus genome was amplified in three overlapping fragments, covering nts 307-870, 715-1302 and 868-483, respectively. 5 µl of reverse-transcribed DNA were amplified. Primers and variable conditions are listed in Table 3. Constant PCR parameters are listed in Table 6.

Table 6: Parameters for PCR using the Phusion™ High Fidelity DNA polymerase

Reagent	End concentration or quantity
Buffer (incl. dNTPs and 1.5mM $MgCl_2$)	1x GC of HF buffer (see Table 3)
Forward primer	0.8 µM
Reverse primer	0.8 µM
SYBR® Green (if appropriate)	1x
DMSO (100%)	See Table 3
Phusion™ High Fidelity DNA polymerase	1 unit
Template	5 µl
RNase/DNase free H_2O	Add up to 25 µl total reaction volume

4.4 Quantification of HBV and HDV by Taqman PCR

Quantitative PCRs were performed using 5 µl of DNA. Primers and variable conditions are listed in Table 2 for HBV and Table 3 for HDV. Constant PCR parameters are listed in Table 7.

Table 7: Parameters for quantitative PCR using the TaqMan Universal PCR master mix

Reagent	End concentration or quantity
TaqMan Universal PCR master mix	1x
Forward primer	See Table 2, Table 3
Reverse primer	See Table 2, Table 3
Probe	See Table 2, Table 3
Template	5 µl
RNase/DNase free H_2O	Add up to 25 µl total reaction volume

5 Agarose gel electrophoresis

Amplified PCR products were separated by size on an agarose gel. Depending on the size of the amplified product, the agarose concentration was varied between 0.7% and 1.5%, with shorter fragments requiring a higher agarose concentration. For a standard gel (approx. 14x12 cm) the appropriate amount of powdered agarose was dissolved in 100 ml of 1x TAE buffer, using a microwave oven. For a big gel (approx. 25x25 cm) 300 ml TAE buffer was used. Then, 10 µl SYBR® Safe for a standard gel or 30 µl for a large gel were added, the warm gel was poured into a casting form and combs were inserted.

After polymerization, the gel was placed in a gel running chamber with sufficient 1xTAE to cover the whole gel and combs were removed. 5µl of amplified DNA was mixed with 1 µl of 5x loading dye and loaded into the resulting pockets. On each gel at least one lane of marker was included. The gel was run at 130V for 40 min for a

normal gel and 60 min for a large gel. After completion, DNA bands were visualized under ultraviolet light (300nm) and pictures were taken with the InGenius Gel documentation System.

6 Agarose gel extraction

In case of unspecific DNA amplification or for cloning purposes 20 µl to 25 µl of PCR product were mixed with 5 µl of 6x Loading Dye and separated on a standard sized agarose gel. Bands were visualised using the Safe Imager™ and excised using a clean scalpel. Purification of DNA was performed using the QIAquick® Gel Extraction Kit according to manufacturer´s instructions. Briefly, the excised gel slice was placed in a 2 ml reaction tube, whose weight had been determined, and weighed. Three volumes of QG buffer were added to one volume of gel, with 100 mg corresponding to 100 µl. The tubes were incubated in a heating block at 50°C until the gel was completely dissolved, (for approximately 10 min). The mixture was vortexed and briefly centrifuged to remove droplets from the lid. Then one gel volume of isopropyl alcohol was added. After mixing, the solution was transferred to a QIAquick spin column (placed in a 2ml collection tube) and centrifuged at 12,000g in a table-top centrifuge for 1 min. The flow through was discarded. The column was washed with 500 µl of GQ buffer and centrifuged at 12,000g for 1 min. Again, the flow through was discarded. The column was then washed with 750 µl PE buffer and centrifuged at 12,000g for 1 min. The flow through was discarded and the column was centrifuged dry at 12,000g for 1min. The DNA was eluted into a 1.5 ml tube by application of 30 µl EB elution buffer to the membrane, incubation for 1 min and centrifugation at 12,000g for 1 min. Eluted DNA was stored at -20°C.

7 PCR Product Purification

Using the Jetquick PCR product Purification Spin kit, PCR products that were not subjected to gel extraction were purified to remove residual primer, non-incorporated nucleotides and PCR reagents, before subjecting them to downstream application such as cloning or sequencing. In short, 20µl of PCR product were mixed with 175µl of H1 buffer, transferred to a Jet Quick Spin® column (placed in a 2ml collection tube) and centrifuged at 12,000g in a table-top centrifuge for 1min. The flow through was discarded. The columns were then washed with 500 µl of H2 buffer and centrifuged at 12,000g for 1 min. The flow through was discarded and the

column was again centrifuged dry at 12,000g for 1min. The DNA was eluted into a 1.5 ml tube by application of 30 µl TE elution buffer (heated to 70°C) to the membrane, incubation for 1 min and centrifugation at 12,000g for 1 min. Eluted DNA was stored at -20°C.

8 Cloning

Cloning was used to identify different viral variants, such as different genotypes or subgenotypes within an individual sample and to acquire larger DNA quantities for downstream applications. Purified PCR products were cloned into a TOPO® vector and transformed into TOP10 *E. coli* electrocompetent cells using the TOPO TA Cloning® Kit or the Zero Blunt® TOPO® PCR Cloning kit.

8.1 TOPO TA Cloning®

During PCR, the Platinum® Taq DNA polymerase adds 5' adenosine (A) overhangs to every PCR product, irrespective of the sequence of the template. This feature is used by the TOPO TA Cloning® kit, which uses a linear pCR®4-TOPO® vector with 3' thymidine (T) overhangs and a covalently bound topoisomerase. Due to the complimentary overhangs, the PCR product is included into the vector backbone, leading to a circular closed plasmid.

For cloning, 4 µl of PCR product were mixed with 1 µl of a 1:10 diluted salt solution and 1 µl of TOPO TA Cloning® vector (included in the kit). The reaction mix was incubated for 5 minutes at room temperature and then put on ice. Electrocompetent One Shot® TOP10 *E. coli* were diluted 1:1 with distilled water. 50 µl of diluted bacteria were pipetted in a 0.1 cm electroporation cuvette (on ice). After adding 4 µl of the ligated vector, the vector/bacteria mix was electroporated at 2.25 kV, 200 Ohm and 25 µF. 250 µl of warm S.O.C. ® (37°C) were directly added after the electroporation and the solution was transferred into a 2 ml reaction tube and incubated at 37°C for 1 hour to allow phenotypic expression of resistance genes. For blue and white screening of the colonies, 40µl of X-Gal was spread onto LB plates containing 30µg/ml Kanamycin. These plates were also incubated at 37°C for 1 hour. After incubation, 50 µl and 150 µl of each cloning reaction were spread on individual LB_{Kan} plates and incubated at 37°C over night.

8.2 Zero Blunt® TOPO® PCR Cloning

PCR amplification using the Phusion™ High Fidelity DNA polymerase results in blunt-ended products. Thus, cloning of these PCR products was performed using the Zero Blunt® TOPO® PCR Cloning kit. Briefly, 4 µl of PCR product were mixed with 1 µl of a 1:4 diluted salt solution and 1 µl of Zero Blunt® TOPO® vector. The reaction mix was incubated for 5 minutes at room temperature and then put on ice. The downstream cloning steps were performed as described in chapter 8.1 TOPO TA Cloning®, p. 42).

8.3 M13 PCR

To confirm the cloning of PCR products into plasmids and in order to sequence the plasmid inserts, bacterial colonies were subjected to M13 PCR ("colony PCR"). As the M13 primer binding sites are located on the plasmid, amplification of the insert is independent of its sequence. Single bacterial colonies were picked as a template from a LB growth plate, using a sterile tooth pick or pipette tip. PCR was performed under the condition described in Table 4. Constant PCR parameters are listed in Table 5.

9 Sequencing

Sanger sequencing, also called dye terminator sequencing, was performed using the BigDye Terminator® v3.1 Cycle Sequencing kit and a mixture of dNTPs and fluorescently labelled, chain terminating dideoxynucleotides ddNTPs (Table 8). The PCR reaction was denatured at 96°C for 1min, followed by 25 cycles of denaturation at 96°C for 10s, annealing at 50°C for 5s and elongation at 60°C for 2min. PCR products were purified using 5 µl 125mM EDTA and 10 mM RNase/DNase free H_2O. After mixing, 60 µl 100% Ethanol was added and the samples were mixed and incubated for 15 min at room temperature in the dark. Samples were centrifuged in a pre-cooled (4°C) centrifuge for 30 min at 3,000 rpm. The solution was then removed. Washing was repeated as described, using 70% of Ethanol. The solution was again removed and samples were dried for 15 min in a vacuum centrifuge and stored at 4°C until further processing. The sequencing reaction was performed on the capillary sequencer ABI PRISM® 3130xl Genetic Analyzer, using 80cm capillaries. For this, samples were heated to 95°C for 5min, supplemented with 10µl Hi-Di™ Formamide and incubated at 95°C for 5min before loading onto the capillary sequencer.

Table 8: Sequencing PCR using the BigDye Terminator® v3.1 Cycle Sequencing kit

Reagent	Concentration or quantity
BigDye Terminator® Mix	1x
TE buffer	1x
Primer forward	0.5µM
Primer reverse	0.5µM
DNA template	10ng (max. 5 µl)
RNase/DNase free H_2O	Add up to 10 µl total reaction volume

10 Phylogenetic analysis

10.1 Electropherogram

Sequence electropherograms obtained from the capillary sequencer were extracted using the SeqScape® software and aligned to a known reference sequence. Furthermore, multiple sequences belonging to the same sample were assembled to larger fragments or complete genome sequences. The electropherograms were visually inspected for inconsistencies. Sequences were then imported into the BioEdit sequence for further processing. Multiple ambiguous nucleotides in SeqScape® throughout the partial sequences or in sequence overlaps indicated mixed infection with multiple HBV variants.

10.2 Sequence alignment

Sequences were aligned by genetic similarities, aiming at minimizing nucleotide or amino acid discrepancies. Gaps are introduced by the alignment program when they increased overall similarity. Pairwise alignments were performed using BioEdit with the Clustal W option and MAFFT v6 with the L-INS-i option, with manual corrections.

Consensus-50 analyses were performed using BioEdit. Sequences of individual subgenotypes were compressed with a threshold of 50% for inclusion in the consensus nucleotide sequence. Consensus-50 subgenotype sequences were used to reconstruct genotype consensus-50 sequences in order to exclude a bias of the differently sized subgenotype sequence sets available.

10.3 Recombination analysis

Recombinant strains can be detected in individual sequences by phylogenetic reconstruction of different overlapping sequence fragments or windows to a

reference dataset. A recombinant strain is detected, when different fragments of the analysed sequence are attributed to different viral variants or (sub)genotypes. Recombination analyses were performed using the Bootscan analysis in the Simplot v3.5.1 software, jpHMM and the RDP, GENECONV and MaxChi options of RDPv.3.44.

10.4 Nucleotide distance calculation

Nucleotide distances between sequences and within or between groups of sequences were calculated using the MEGA software and the neighbour-joining method of the Kimura 2-parameter model.

10.5 Phylogenetic and phylogeographic reconstruction

Phylogenetic trees are inferred based on the analysed sequence alignment. The tree can be unrooted or rooted. Sequences are represented on branches that are connected by nodes. Sequences clustering on one node are in the same phylogenetic group. Phylogenetic groups are, e.g. referred to as genotypes and subgenotypes (hepatitis B virus) or clades (hepatitis D virus). In addition to the branching pattern, the length of the branches and the horizontal distances between sequences are important. Statistical methods are used to determine the branch length and genetic relatedness between sequences. To assess the reliability of the inferred branching pattern, bootstrap procedures are performed, resampling sequence subsets and comparing these with the originally calculated trees.

Several different methods are available for reconstructing phylogenetic trees. The most common construction method is the neighbour-joining method. For analyses, the MEGA software was used with the neighbour-joining method of the Kimura 2-parameter model and with 1,000 bootstrap replicates.

In the case of more complex phylogenies or for phylogeographic reconstruction computationally extensive Bayesian Markov chain Monte Carlo (MCMC) analyses are applied. For analyses the BEAST v.1.6.2 software package was used. Briefly, BEAST averages phylogenetic trees over tree space to weigh it proportional to its posterior probability. Phylogenies are inferred using strict or relaxed molecular clock models to infer rooted, time-measured phylogenies. In BEAST analyses parameterized runs were repeated to reach an effective sample size greater than 200. The most applicable model was selected by calculating Bayes Factors of the different parameterized runs.

Phylogeographic analyses were performed using BEAST by including geographic state parameters. Visualisation was performed using Google Earth v6.2.1 and figures were created using Microsoft Office Powerpoint.

10.6 Median-joining network

Median-joining networks (MJN) were constructed using Network v4.610 and visualized using Network Publisher. Briefly, MJN were constructed starting from minimum spanning trees that were combined within a single network. Aiming at parsimony, consensus sequences of 3 mutually close sequences were subsequently added at a time. These so called median vectors can biologically be interpreted as extinct ancestral sequences. This median operation then resulted in the most optimal network, reflecting the shortest possible distances between the individual nodes (20).

Chapter IV:Results and Discussion

Part 1: Hepatitis B virus: the genotype E puzzle

This manuscript has been published as a Review article:

Andernach IE, Hübschen JM, Muller CP. Hepatitis B virus: the genotype E puzzle. Rev Med Virol. 2009 Jul;19(4):231-40.

I.E. Andernach contributed significantly to the design of the experimental part and structure of the manuscript and was in charge of literature research, data collection, data analysis and writing of the manuscript.

Hepatitis B virus (HBV) is highly endemic throughout sub-Saharan Africa. One of the two genotypes A and E dominates in most countries. With several subgenotypes and variants, genotype A is more diverse in Africa (4.00%) than in the rest of the world (2.96%), suggesting an African origin and a long history on the continent. Despite the African slave trade, genotype E has only sporadically been found within the Americas, indicating that this genotype was introduced only during the past 200 years into the general African population. A short history for this genotype in Africa is also supported by its conspicuously low genetic diversity (1.75%), which contrasts, however, with its excessively high HBsAg prevalence and its extensive spread throughout the vast West-African genotype E crescent. We discuss the spread and routes of transmission of genotype E and suggest that the distribution and current high prevalence levels of HBV (genotype E) in Africa are the result of the extensive use of unsafe needles, potentially solving the current African genotype E puzzle and shedding new light on the high HBV prevalence in Africa.

1 Prevalence of hepatitis B virus in Africa

Despite efficient vaccines, HBV infections continue to be a public health problem. It is estimated that more than a third of the world population is or has been infected with HBV at some time during their lives (306). More than 350 million people are chronic carriers of the virus (309), of which at least 65 million live in Africa (147, 197).

Worldwide, HBV infections account for 1 million deaths per year (309), 250 000 in Africa alone (147). Acute hepatitis B infection is cleared in more than 90% of adult cases (306). During early childhood, acute infection is mild or inapparent, but is associated with a risk of chronic disease that can be as high as 90% in infants (82, 110). Chronic carrier status is relatively low throughout most of Western Europe, whereas in South-East Asia, the Indian sub-continent and throughout most of sub-Saharan Africa the HBsAg prevalence exceeds 10% (156). In high risk populations it can be much higher, e.g. 55% in medical students in Mali (197) or more than 90% in HIV carriers in Cameroon (197). The prevalence of anti-HBc antibodies is even higher and reaches about 85% in Western Africa and 65-85% in East Africa (147), indicating that in many sub-Saharan countries most people have at some time been infected with HBV.

2 HBV genotypes

By accumulating mutations during its natural history, HBV has evolved into at least eight recognised genotypes A–H (145, 146, 209-211, 217, 252, 270) plus a potential new genotype (tentatively designated genotype I) proposed for strains mainly found in Laos (125, 219), but also in Vietnam (106, 294). With the exception of genotypes E, G and H, genotypes are divided into subgenotypes with more or less distinct geographic distributions (75, 107, 117, 119, 120, 133, 142, 145, 150, 152, 202, 209, 221, 223, 273, 274). HBV was also found in nonhuman primates, including chimpanzees (ChHBV), orangutans (OuHBV) and gibbons (GiHBV) (101, 179, 209, 249), each with its own distinct genotype.

3 Genotype A in Africa

The majority of non-E genotypes reported from Africa belong to genotype A (Figure 10, panel 1). Although only relatively few HBV/A sequences are available from African countries, except for e.g. South Africa (142, 224), Malawi, Tanzania (109) and Somalia (209), it seems that subgenotype A1 is the dominant genotype in countries situated along the Eastern coast from South Africa to the Horn of Africa in the North and Rwanda in the West (117, 147). In contrast, in West-Africa, HBV/A strains are relatively rare and belong to different subgenotypes (Figure 10, panel 2). We have found three new variants in West-Africa, tentatively designated subgenotypes A3 (originally A'') (197), A4 and A5 (221), and a potential new subgenotype in Rwanda.

Subgenotype A3 was originally identified in Cameroon (197), but was also found e.g. in The Gambia (107), Gabon (180) and Nigeria (221). A4 has been reported from The Gambia and Mali (197) and A5 from Nigeria (221). A2 has rarely been found, but mainly in South Africa (224). Nevertheless the genetic diversity of A2 is higher in Africa (1.97%) than in Europe (1.38% mean genetic diversity over the complete genome), where it is highly prevalent. It has therefore been suggested that this subgenotype originated from Southern Africa and was brought to Europe by sailors in the fifteenth century (107). As a result of the different subgenotypes, the genetic diversity of genotype A is higher in Africa (mean diversity of 4%) than anywhere else in the world (2.96%) (Table 9). Hence it has been speculated that genotype A may have initially evolved in Africa, from where it has spread to the other continents, giving genotype A many centuries of evolutionary history in Africa.

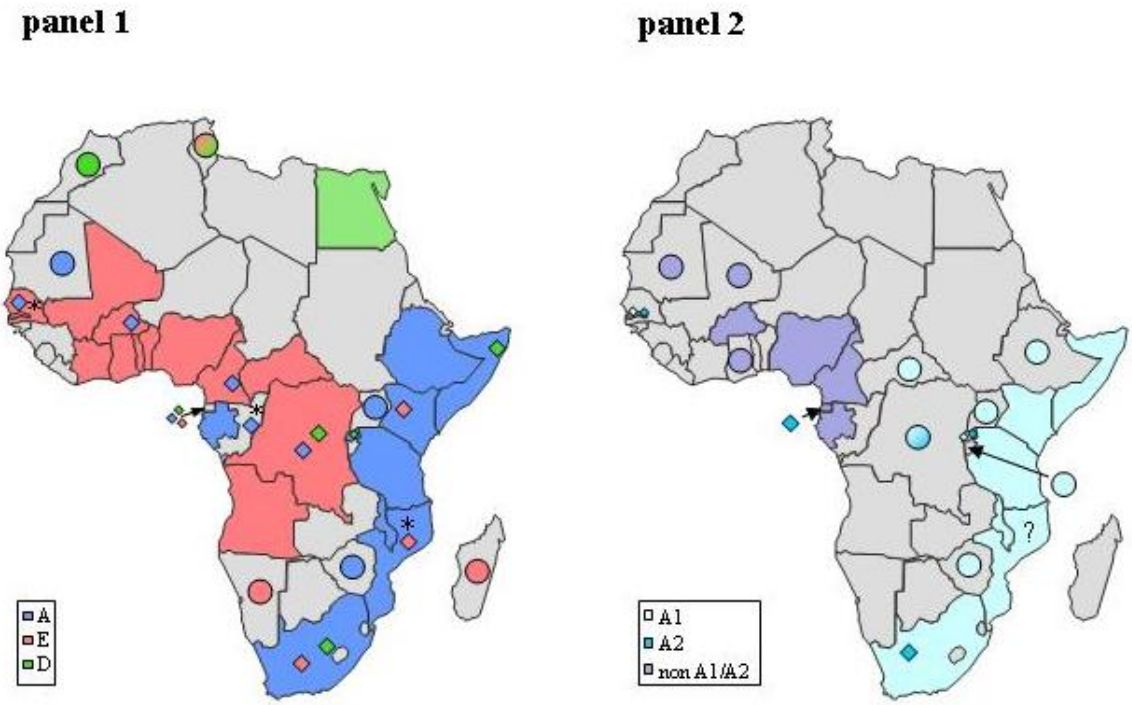

Figure 10: Distribution of HBV genotypes (panel 1) and genotype A subgenotypes (panel 2) in Africa.
Countries with overall less than 10 strains (panel 1) or less than 10 strains of genotype A (panel B) available are marked with a coloured dot. Countries with full colour have >90% dominance of one (sub-)genotype or a 60-90% predominance of one (sub-)genotype with minority (sub-) genotypes shown as diamonds. For countries marked by asterisks, no sequences were available and shading is modified from reference (147). The A subgenotypes in Mozambique are unknown. Sequences obtained from GenBank and unpublished data were included.

4 Genotype E in Africa

Genotype E was first described in 1992 (211), but its ubiquitous spread and high prevalence throughout the 'African genotype E crescent' has only recently been recognised. In large studies including hundreds of sequences from The Gambia (79), Mali (197), Burkina Faso (197), Ghana (48, 197), Togo (197), Benin (197), Nigeria, Central African Republic (28) and Democratic Republic of Congo (197) this genotype was found to be by far the most prevalent in these countries (including unpublished data). Genotype E also seems to be the main genotype in Senegal (301), Ivory Coast (278), Angola and Namibia (147).

Interestingly, in Cameroon the overall prevalence of genotype A and E is similar, but large differences between cohorts have been reported (152, 197, 221). In a large HIV cohort, 91.4% were HBsAg positive and 82 and 18% carried genotype A3 and E, respectively (197). In children, genotype E prevailed (78%) (197). Bantus and the relatively isolated Pygmies showed an equal distribution of genotype A3 and E (152). Genotype E strains from this country did not form a separate cluster, but were interspersed among sequences from other countries and were as diverse as the latter. As genotypes E and A co-circulate in Cameroon, mixed infections are frequent (37%) and several A/E recombinations were reported (152, 221). In Europe genotype E has been found in sporadic cases, mostly in African immigrants to countries of the Mediterranean basin (91, 293).

Table 9: Mean genetic diversity of HBV genotypes A and E, based on complete genome and S-fragment sequences, available on GenBank and unpublished data.

	N	Mean intra-group diversity (%)			
Genotype	FL (S)	Worldwide	Africa	Non-African countries	Europe
A	131 (202)	4.00 (2.10)	4.00 (2.09)	2.96 (1.76)	/
A1	55 (89)	2.85 (1.36)	2.74 (1.36)	2.28 (0.97)	/
A2	49 (63)	1.24 (1.25)	1.97 (1.59)	1.17 (1.06)	1.38 (1.41)
non A1/A2	24 (50)	n.a. / (1.85)	3.51 (1.85)	/	/
E	69 (610)	1.75 (0.79)	1.75 (0.79)	/	/

The genetic diversity of S-fragment sequences is shown in brackets. FL=full length sequence, S=S-gene sequence, n.a.=not applicable.

Despite its wide geographic spread and high prevalence, genotype E viruses in Africa reveal a surprisingly low genetic diversity. When all currently published and some unpublished complete genotype E sequences are included (n=69), the mean diversity over the whole genome is 1.75%, in comparison to 4% diversity for all African genotype A sequences (n=66). The same analysis of S-fragment sequences (n=610) shows a lower mean genetic diversity of 0.79%, compared to 2.09% for African HBV/A sequences (n=167) (Table 9). Careful analysis of a limited number of complete genomes seemed to suggest a higher diversity in the Northern part of the genotype E crescent (from Benin to Cameroon) and a separate cluster with a lower genetic diversity and a common node in Democratic Republic of Congo, Angola and Namibia, the Southern half of the genotype E crescent (115, 147). This may be an indication that this genotype emerged in the North before spreading to the South. Overall, the low genetic diversity is indicative of a short evolutionary history and a more recent introduction of genotype E than genotype A into the general African population. It has been suggested that it would take about 200 years to develop the current mean diversity of genotype E (197).

5 HBV genotype E and A in the Americas

The spread of pathogens is closely linked to their host, e.g. phylogenetic analysis has shown that the human T-lymphotropic virus 1 was brought from West-Africa to the Americas after the 15th century (296). Similarly, African slaves that were forcemigrated to the Americas would have disseminated African HBV strains in the New World.

Nevertheless, HBV/E is only sporadically found in the Americas, despite the forced migration of these slaves to the New World between the 16th and 19th centuries. Among more than 1600 sequences from 26 studies (and GenBank, December 2008), 47.4% were of genotype A, 31.8% of genotype F and 12.6% of genotype D. Some sequences belonged to genotypes H (2.9%) and C (3.2%) and only sporadic sequences to genotypes G (0.8%), B (0.7%) or E (0.2%, Argentina, Brazil)(71, 74, 184-186, 193, 196, 203, 205, 238, 240, 267, 298).

In most South-American countries the indigenous genotype F prevails (Figure 11), while in Brazil, with the largest community of African descendants (44.7%) in South- and Central America (65), a high prevalence of genotype A has been reported (10). Even in an isolated Afro-Brazilian community almost all strains were assigned to

genotype A1 (196), otherwise found typically in Eastern Africa. This can be explained by the transatlantic slave trade, as Brazil was the largest single recipient of slaves from West and West Central Africa and later increasingly from East Africa (e.g. Mozambique)(132). In Afro-Venezuelans, half showed genotype F and the other half genotype A2 (239). All HBV/A sequences from the Americas that could be subgenotyped, were exclusively assigned to subgenotypes A1 (Brazil, Argentina) and A2 (Venezuelans including Afro-Venezuelans, Brazil, Argentina, Mexico, Nicaragua). However, more complete genome sequences from Latin America are required to understand their relationship with current African sequences.

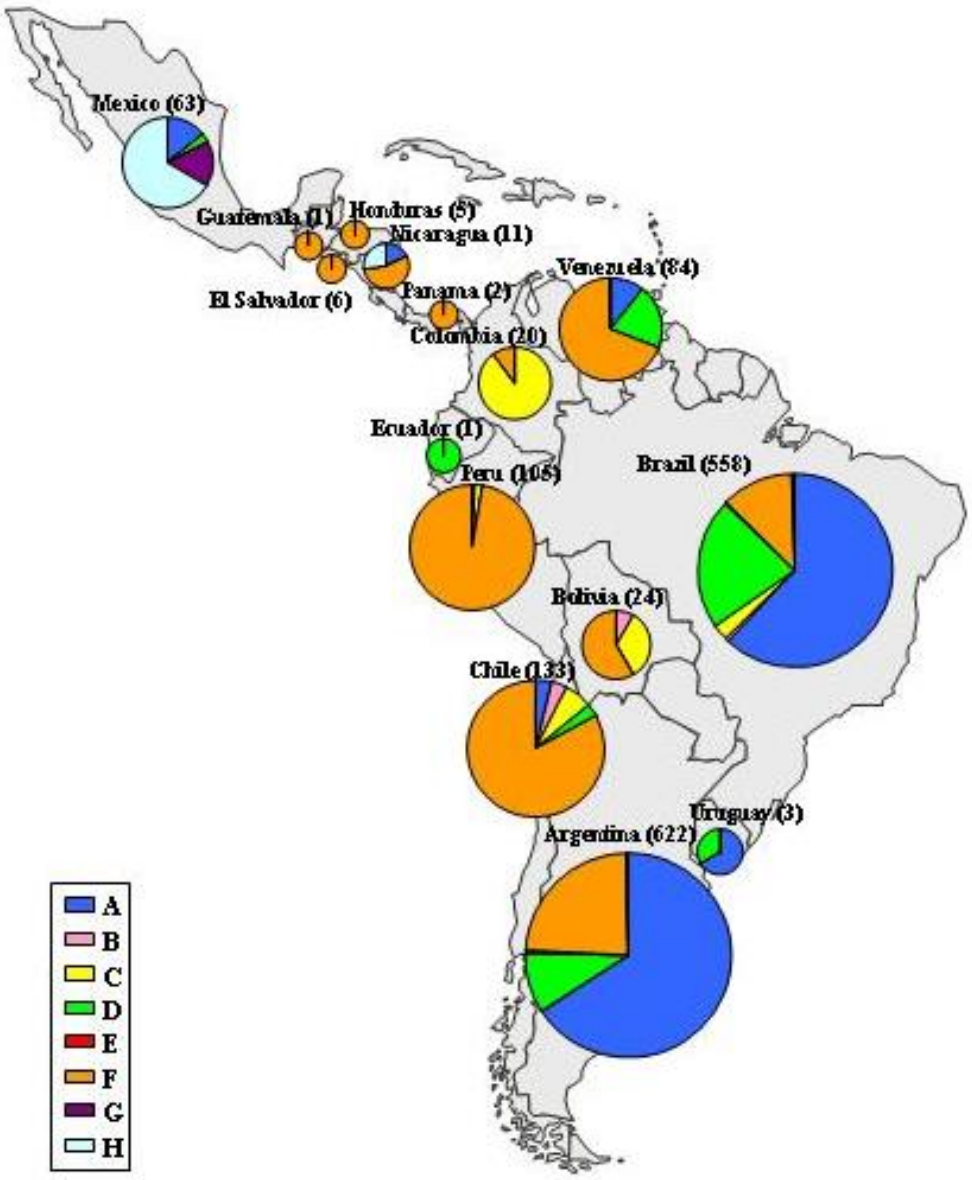

Figure 11: Geographic distribution of HBV genotypes in South and Central America. Single HBV/E sequences are not visible in the figure. The number of strains is shown in brackets together with the country name.

6 The HBV/E puzzle

6.1 The role of the slave trade

The conspicuous absence of a significant number of genotype E strains from the Americas and the dominance of HBV/A in Afro-Americans strongly suggest that slaves were not carriers of genotype E, indicating that genotype E was rare when and where the slaves were rounded up in Africa. Slaves originated mostly from three main regions in West-Central Africa: the Bight of Benin, the Bight of Nigeria and the Gold Coast in Ghana (194). Thus, genotype E must have been introduced into the general West African population only after the slave trade had come to an end, which was, throughout most of the Americas, only by the turn of the 18th century (194). This would be compatible with the 200 years for the current mean genetic diversity of genotype E to develop. Alternative explanations would postulate that conditions in the Americas were not conducive to the continued transmission of HBV/E, but would fail to explain the distribution of African HBV/A strains in the Americas.

6.2 No known source of genotype E

Genotype E has been reported to be more closely related to non-human HBV isolates than the other genotypes, except genotype F (179). A HBV/E virus (GenBank accession AB032431) has been found once in a chimpanzee, but the direction of transmission was not established (280). Cross-species transmission from humans to baboons (139) and gibbons (19, 258) was shown experimentally, but there is no evidence of systematic natural transmission of HBV from non-human primates to humans (249).

An introduction of HBV/E from the closest known non-primate (woodchuck, 53.5% mean genetic distance) and primate animal reservoir (chimpanzees, 10.1%) would have taken more than a thousand years to cover the mean genetic distance across the whole genome (83). Also the evolution of E from the closest known genotype (D), or from the African genotype A would have taken many centuries to bridge the 8.11 or 10.73% mean genetic distance. As the evidence seems to corroborate an introduction of genotype E into humans within the last 200 years, it seems unlikely that known nonhuman primate genotypes are the source of a recent introduction of HBV/E into the general population, although primates would be the most probable source of such a putative ancestor strain.

6.3 Cross-protection between genotype A and E

The much higher genetic diversity of HBV/A than of HBV/E suggests that genotype A was widespread in Africa long before genotype E was introduced. If natural routes of transmission were as effective for HBV/A as for HBV/E, genotype E had to overcome widespread pre-existing immunity to HBV/A viruses. Antibodies against important epitopes of the HBsAg are critical for protection and are mainly directed against the major antigenic determinant 'a' (amino acid 124–147) (330). Both genotypes E and A show consistent amino acid differences in the a-determinant (Table 10), which may reduce cross-protection between genotype E and A strains, although amino acid position 145, which is often associated with vaccine escape mutations against genotype A based HBsAg vaccines (14, 330), is not one of them. A potential displacement of HBV/A by genotype E seems therefore not plausible.

Table 10: Characteristic amino acid changes in genotype E and A strains found across the major "a" determinant present in HBV vaccines.

		Amino acid position in HBsAg					
Genotype	N	122	127	131	140	143	159
E (679)	598	R	L	T	S	S	G
A1 (70)	40	K	P	N	T	T	A
	21	R	.	.	.	.	.
A2 (30)	22	K	P	N	T	T	A
	7	R	.	.	.	.	.
non A1/A2 (42)	28	R	P	N	T	T	A
	8	K	.	.	.	.	.

Sequences harbouring mutations with a prevalence < 10% per genotype are not shown. Total numbers of sequences of each genotype or subgenotype are shown in brackets. Position 122 co-determines the serotype.

6.4 Vertical transmission

HBV infections tend to be mild or asymptomatic in children, but 90% of infants and 30–50% of toddlers become chronic carriers, in contrast to 5–10% of adults (306). Thus the age of infection is critical for the spread of HBV, but, in contrast to Asia, perinatal infection or infection in early infancy is estimated to be as low as 1–8% in sub-Saharan Africa (48, 140, 246). Nevertheless, early childhood transmission is thought to be the most important infectious route with most children being infected by the age of 5 years (48). The capacity to produce HBeAg and seroconversion to anti-HBe antibody is thought to play an important role in the transmission and outcome of perinatally acquired HBV infections (48, 140). The risk of perinatal infection increases if the mother is HBeAg positive and 80% of their infants become chronic carriers (140). Seroconversion to anti-HBe in Ghana and Benin, where HBV/E is predominant, occurs much earlier in life (<16 years of age) than in Asia (48, 88) and thus women of child-bearing age are less likely to infect their children. The prevalence of anti-HBe seems to be similar for genotypes E and A (91). Thus, vertical and early horizontal household transmission alone could hardly account for the high seroprevalence that would have developed throughout the HBV genotype E crescent within the estimated last 200 years.

6.5 Horizontal transmission

The recent introduction of HBV/E seems difficult to reconcile with the excessive HBV prevalence throughout the genotype E crescent, unless a very efficient transmission mechanism has caused its explosive spread.

During the last 200 years, numerous injection mass-campaigns were performed, e.g. in French and Belgian colonies, where nowadays genotype E largely prevails, against yaws in the 1920s (77, 266) or sleeping sickness, where, for instance, from 1917 to 1919 only six syringes were used to treat 90 000 individuals (89). Also smallpox campaigns in West Africa, e.g. in the French colonies where at least 35 000 immunisations were documented between 1893 and 1910 (111) and the pervasive use of injectable antibiotics have resulted in frequent and unsafe inoculations (77). In Egypt, the widespread and efficient transmission of hepatitis C virus HCV was recently linked to unsafe injections during mass-treatment of schistosomiasis until the 1980s (87). The more stable HBV is estimated to be 10 and 20 times more transmissible via unsafe injections than HCV and HIV, respectively (87, 266). According to some studies at least 13 million HBV infections per year may be caused by unsafe injections in the

developing world (266). In addition to virus transmission as a result of poor iatrogenic practices, permissive sexual behaviour and traditional tribal practices during adulthood also contribute to the horizontal spread of HBV. However, conversion to chronic carrier status would be infrequent. In this context, insect vectors of infection should also be revisited using modern sensitive detection techniques. So far mainly bedbugs and kissing bugs have been implicated in the spread of HBV, as the persistence of HBsAg and DNA in these organisms and their excrements was shown (31, 264). Also mosquitoes could potentially be infective, but only when changing hosts after an interrupted blood-meal (32). Interestingly, the geographical distribution of HBV/E matches largely that of the tsetse fly which has been considered a potential vector of HCV (237). Whether a combination of once-in-a-generation early-life vertical infection associated with a high risk of chronicity and rampant infections during adult-life with a lower risk of chronicity could explain the explosive spread of HBV/E, requires sophisticated modelling combined with demographic and epidemiological assumptions.

6.6 Conclusions

The genetic diversity of HBV/A suggests that this genotype has a longer evolutionary history in Africa than genotype E. This is also confirmed by observations that HBV/A is the most important genotype in Afro-Americans. The absence of genotype E in the descendants of the African slaves and the low genetic diversity suggest that this genotype was introduced during the past 200 years, after the slave trade had come to an end. Although the ancestor of HBV/E is not known, evidence points towards a secluded tribe or non-human primates as a historical reservoir in Western Africa. The excessively high prevalence of HBV/E throughout the genotype E crescent can only be explained by an explosive spread of the virus. Thus, routes of transmission need to be revisited. Although a competitive advantage of HBV/E in comparison to HBV/A cannot be fully excluded, there is no direct evidence that genotype E is more infectious. If routes of transmission had been the same for both genotypes, it could have been expected that, before the introduction of HBV/E, genotype A would be as prevalent throughout Africa as genotype E is today. In this case genotypes E and A would have had co-circulated for many years before genotype E had overcome widespread pre-existing immunity against HBV/A to eventually replace it. As this is not very likely within the short time frame, two of the above assumptions must be revised: (i) The prevalence of genotype A and thus of HBV was traditionally much lower in

Africa than the HBV prevalence is today and HBV/A only occurred in relatively secluded areas. This is supported by the low prevalence of genotype A in the HBV/E crescent; the rare recombinations between genotypes A and E also suggest that the clash between the two genotypes historically never occurred. (ii) The explosive spread of HBV/E must have been the result of a new route of transmission, 'revolutionary' at that time. Much evidence points towards historical mass-campaigns with unsafe injection needles as the main culprits of the current high prevalence levels of HBV and in particular genotype E throughout most of sub-Saharan Africa. This seems to be the one and most obvious way to reconcile all genotype E contradictions and to explain the current African genotype E puzzle. It also sheds new light on the high HBV prevalence that is seen today in Africa.

Part 2: Hepatitis B virus genotype E variability in Africa

This manuscript has been published as a Review article:

Hübschen JM, Andernach IE, Muller CP. Hepatitis B virus genotype E variability in Africa. J Clin Virol. 2008 Dec;43(4):376-80.

I.E. Andernach contributed to the data collection, data analysis and preparation of the manuscript.

Despite the availability of efficient vaccines, hepatitis B virus (HBV) is estimated to cause chronic infections in more than 350 million people worldwide and death in one million per year (309). Nearly 20% of the chronic carriers live in Africa with excessively high prevalence rates reported especially from the sub-Saharan regions (147, 152, 180, 231). In Africa, the spread and evolution of HBV genotypes are still poorly understood. While genotype D seems to be the most prevalent genotype in the northern part (147), the highly diverse genotype A is the dominant genotype in the southern and eastern regions of the African continent (107, 142, 147, 150, 275). Genotype E predominates throughout a vast crescent spanning some 6000 km from Senegal to Namibia (147, 177, 197) and extending to the Central African Republic in the East (27, 197). With about 70–140 million chronic carriers in this region, genotype E may well be the most important genotype worldwide (197). Here we review our current understanding of this genotype reanalysing all currently available sequences including many unpublished sequences.

1 Geographical distribution of genotype E

After sequence analysis of more than 400 HBV strains, we found that genotype E was by far the most prevalent genotype in Mali, Burkina Faso, Nigeria, Togo, Benin, Central African Republic and The Democratic Republic of the Congo (27, 197, 215, 221). In Cameroon, either genotype A or E were described to be the most frequent, depending on the cohorts studied (152, 197, 221). As a result of the co-circulation of

these two genotypes, recombinations and a high number of mixed infections were detected (221). Genotype E was also reported from other West African countries like Senegal (301), Ivory Coast (278), Ghana (50), The Gambia (80), Angola and Namibia (147). In addition, a few strains were found further east like in Mozambique (68) and Madagascar (149) and Tunisia (16) in the North. Genotype E has not been found outside of Africa, except for a few rare cases mostly in individuals with an African background. In Europe, sporadic genotype E strains were reported from, e.g. France (91, 105), Italy (226), Spain (293), Belgium (170) and The Netherlands (297) and in the Americas from Argentina (184), Brazil (267) and the US (59, 134).

2 Phylogenetic analysis of African HBV genotype E strains

Despite the wide geographic distribution in Africa and the high prevalence throughout large parts of Africa, genotype E has a very low genetic diversity (147, 197, 221). For the purpose of this review all currently available sequences covering the S gene and the complete genome were aligned in order to identify potential genetic clusters that may help to understand the spread of genotype E strains. For the analysis of the S gene, 596 strains including 163 unpublished sequences from 16 countries were available (Table 11). The complete genome analysis was based on 68 sequences (all published) from nine different countries (Table 11). The phylogenetic tree based on the S gene (Figure 12) showed that except for a few very small groups of identical or very similar strains, no clearly distinct clusters of sequences could be detected. The complete genome analysis, however, revealed that the sequences of the Democratic Republic of the Congo (DRC) form an own cluster, supported by a bootstrap value of 99, together with sequences from Angola and Namibia (bootstrap value 93, Figure 13). The majority of the strains from the Central African Republic form a cluster of their own which is however not supported by a high bootstrap value (below 50). The rest of the Central African sequences group with a strain from Cameroon (AB194948). The strains from Ghana form several clusters most of which are not supported by high bootstrap values (Figure 13). Thus, except for the strains from DRC, Angola and Namibia, it seems difficult to recognize geographical clusters of HBV genotype E in Africa.

Table 11: HBV genotype E strains available for phylogenetic analysis

Country	S fragment sequences	Complete genome sequences	Genotype E/other genotypes	Mean (maximal) genetic distance
Angola	8	2	100% (8/8)	0.85%
Benin	17	4	100% (30/30)	1.64% (2.12%)
Burkina Faso	41	0	76% (47/62)	
Cameroon	29	2	64% (74/116)	2.34%
Central African Republic	55	28	94% (62/66)	1.37% (2.73%)
Democratic Republic of the Congo	7	4	64% (7/11)	0.18% (0.25%)
Equatorial Guinea	3	0	21% (3/14)	
Ghana	160	14	96% (307/319	1.92% (3.18%)
Ivory Coast	10	2	92% (12/13)	1.35%
Madagascar	2	1	100% (2/2)	
Mali	20	0	88% (37/42)	
Namibia	6	6	100% (6/6)	0.82% (1.13%)
Nigeria	142	0	96% (251/262)	
Sierra Leone	2	0	100% (2/2)	
The Gambia	70	0	90% (192/214)	
Togo	22	0	100% (22/22)	
Unattributed	2	5		
∑=16	∑=596	∑=68		1.71% (3.54%)

Genetic distances are calculated over the complete genome sequences within each country.

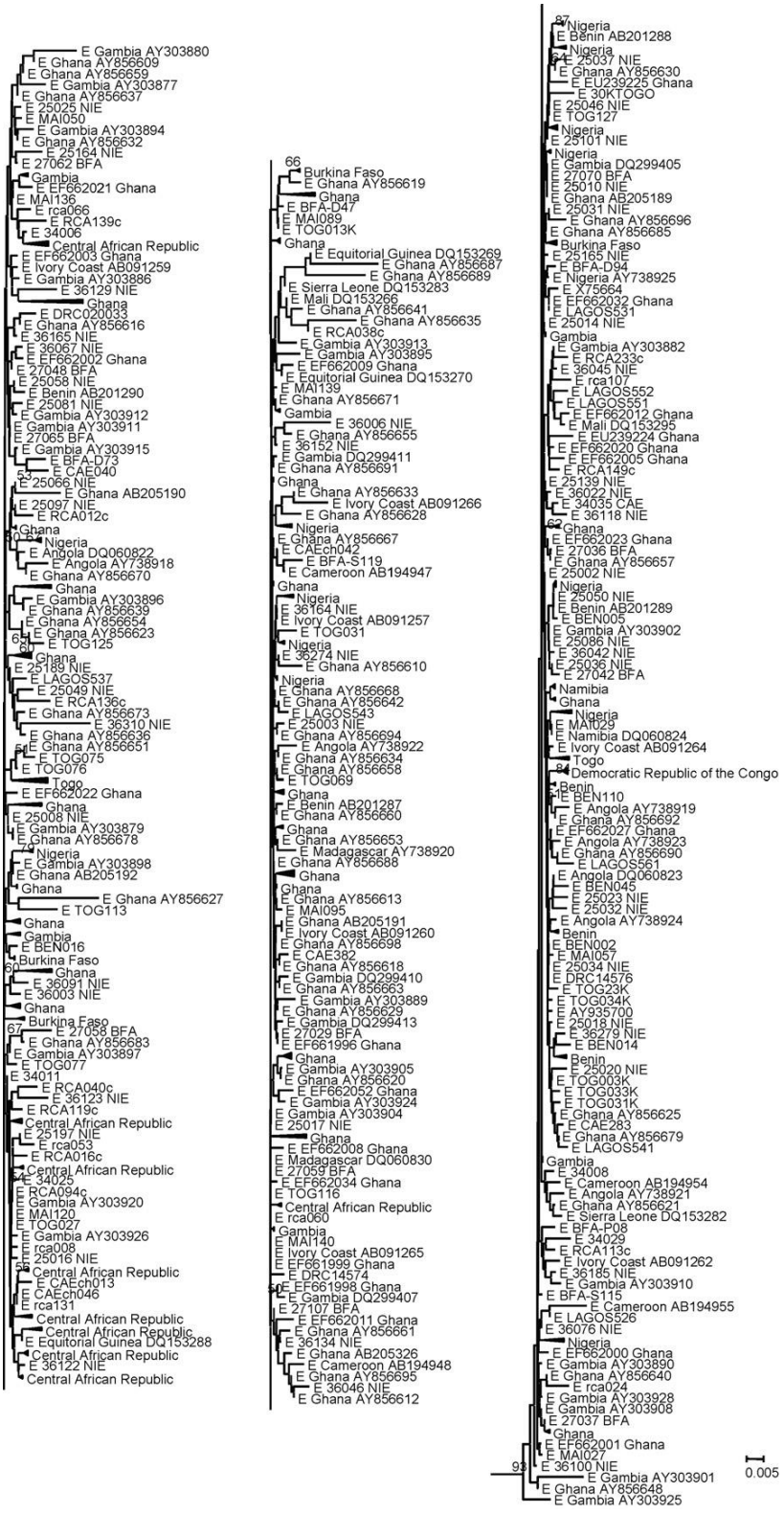

Figure 12: Subdivided genotype E cluster of the phylogenetic tree based on 681 nt of the S gene using all genotype E strains available on GenBank and some unpublished sequences.

Alignments were generated with SeqScape® v2.5 and BioEdit. Phylogenetic analysis was performed using MEGA version 3.1. Only one example of strains from the same country on the same branch is displayed and groups comprising strains from the same country are compressed to reduce the size of the figure. Only bootstrap values above 49 are shown.

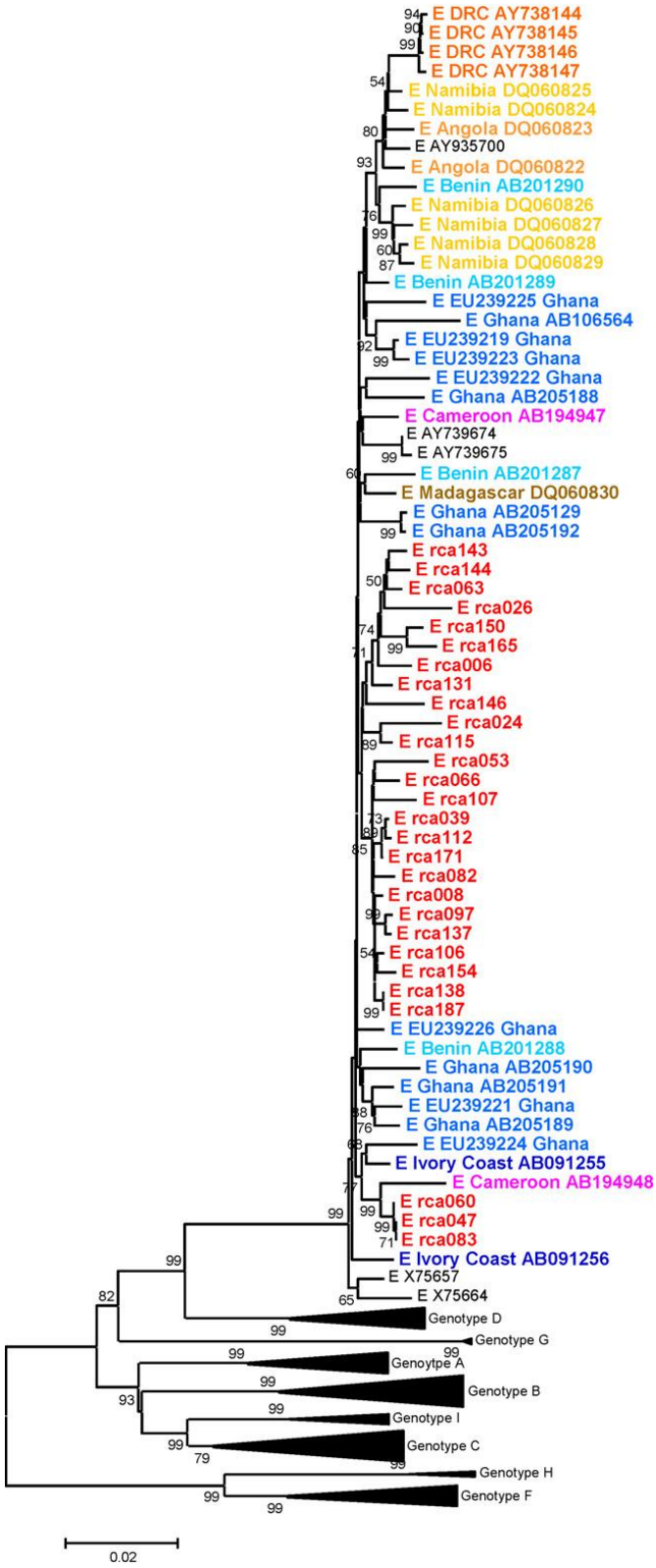

Figure 13: Phylogenetic analysis of complete genome sequences of HBV genotype E using the neighbor-joining method.
Strains from the same country are coded with the same colour. Only bootstrap values above 49 are shown.

3 Genetic distances among HBV genotype E strains

A low mean genetic diversity of about 1.5% for the preS/S and preC/C regions and of 1.73% intragenotype nucleotide divergence over the complete genome of HBV genotype E strains was reported in earlier studies (147, 197, 221). The mean genetic diversity among all currently available genotype E sequences of the S gene and the complete genome was now calculated to be 0.73% and 1.71%, respectively (Kimura 2-parameter method of MEGA version 3.1) (151). The two most distant strains (AB106564 from Ghana and AB194948 from Cameroon or rca026 from the Central African Republic) are separated by a genetic distance of 3.54% over the complete genome. The maximum and mean genetic distances for the different countries are shown in Table 11. The sequences from DRC show a relatively low diversity, indicating that they may have developed only recently from a common ancestor. The situation in Namibia and Angola is similar. The sequences from Benin and Ghana on the other hand show a comparatively high diversity suggesting that genotype E has been endemic in these countries for longer than in the more Southern countries. The sequences from the Central African Republic seem to have a somewhat intermediate diversity. These conclusions need to be verified on a larger number of complete genome sequences, as currently only rare sequences from mostly single studies are available for these countries (Table 11).

4 Recombinations and mixed infections

Except for our own studies, few studies on HBV genotypes in Africa were designed to detect mixed infections or recombinants. As a result of the co-circulation of genotype A and E strains, both recombinations and mixed infections were mainly reported from Cameroon. Mixed infections with genotypes A and E have been detected in more than one third of patients (15/41) from Cameroon and in 4 out of 49 patients (8%) from Nigeria (221). One or possibly two genotype A/E recombinants were described from Cameroon (152, 197, 221); in addition another A/E (or possibly A/D) recombination with a small genotype G insert was reported in a strain from Nigeria (221) and a D/E recombinant from the Central African Republic (27). The first recombinant was a genotype A strain with about 200 nucleotides from the polymerase gene of genotype E (152), the possible A/E recombinant from Cameroon exhibited genotype A in the preS/S part, but genotype E in the preC/C region (197, 221). The Central African Republic recombinant was a genotype E strain

except for a genotype D signature in the C gene (27); the A/E recombination in a strain from Nigeria occurred in the preC/C region (221). For the purpose of this review we checked all published S gene and complete genome sequences of genotypes A and E from Cameroon and all complete genome sequences of genotype E strains for recombinations (SimPlot v3.5.1) (173), but no additional A/E recombinants were found.

5 Discussion and conclusions

Phylogenetic analysis of the nearly 600 S gene sequences revealed no clearly distinct geographic or genetic clusters of strains. However, the phylogenetic tree generated with complete genome sequences showed that the sequences from the Southern end of the genotype E crescent (South-West Africa and DRC) cluster together, as observed before (147), and may be derived from a common source or ancestor. In the North, the West African strains, especially the ones from Ghana and Benin revealed a relatively high diversity and several smaller lineages, indicating that genotype E has a longer natural history in this region. Together, the higher diversity in the North and the lower diversity in the South, suggest that the virus may have emerged in the Northern part of the genotype E crescent before spreading to the South. Our analysis confirmed the South-West/Central African cluster (147), and showed that the 28 new sequences from the Central African Republic (27) included here did not fall into this cluster, but formed at least two distinct lineages, one clustering separately, the other one with a strain from Cameroon.

As genotype E is essentially absent from the Americas despite the Afro-American slave trade until the 19th century, these strains must have been relatively rare when the slaves were rounded up in West Africa. Thus, genotype E strains may have been introduced into the general African population only within the past 200 years. This would be compatible with the conspicuously low genetic diversity of genotype E viruses. However, the origin of these viruses remains obscure. Even the closest known HBV viruses either in humans or non-human primates could not bridge the genetic distance within 200 years. Moreover, how the virus may have spread throughout the vast expanses of the genotype E crescent cannot be explained without revisiting transmission routes. A thorough investigation of possible ways of transmission and spread of genotype E in Africa is warranted to prevent its further dissemination.

Part 3: Slave Trade and Hepatitis B Virus Genotypes and Subgenotypes in Haiti and Africa

This manuscript has been published as:

Andernach IE, Nolte C, Pape JW, Muller CP. Slave Trade and Hepatitis B Virus Genotypes and Subgenotypes in Haiti and Africa. Emerg Infect Dis. 2009;15(8):1222-8.

I.E. Andernach contributed significantly to the study design and experimental work and was in charge of experimental setup, training of the visiting scientist, data analysis and writing of the manuscript.

In sub-Saharan Africa, genotypes E and A predominate. East of the E/A divide (117), subgenotype A1 is dominant in countries along the eastern coast from South Africa to the Horn of Africa (148). Although genotype A has been found on every continent, its genetic diversity is higher in Africa (4% over the complete genome) than in the rest of the world (3%). Therefore, some researchers have suggested that genotype A has emerged in Africa (107) and, after a long evolution, has been introduced to other continents. However, despite the high genetic diversity of HBV/A in West Africa, this genotype is rare there. In contrast, genotype E is found almost exclusively throughout the vast expanses of a crescent from Senegal in the west (301) to the Central African Republic in the east (28) and Namibia in the south (147). In comparison to HBV/A, the conspicuously low genetic diversity of HBV/E suggests its short natural history in Africa (197) and relatively recent introduction into the general population there (197). However, the recent presence of HBV/E in Africa contrasts sharply with its current high prevalence and extensive geographic distribution there. The wide spread of genotype E also seems difficult to reconcile with a long natural history of genotype A in Africa (197).

In Haiti, where >90% of the population descends directly from African slaves (64), we investigated the phylogeny of HBV to learn which genotypes may have been

prevalent in Africa several centuries ago. The conspicuous absence of genotype E in Haiti suggests recent and rapid spread of genotype E in Africa during the past 200 years, probably as the result of public health interventions.

1 Clinical Samples

Serum samples were collected in 2006 after informed consent as part of a national survey to evaluate prevalences of human immunodeficiency virus infection, hepatitis B, and serologic syphilis among pregnant women at their first prenatal medical visit in 19 clinics throughout Haiti. Women were tested for hepatitis B surface antigen (HBsAg) by using the Murex HBsAg Kit. DNA was extracted from HBsAg–positive samples by using the DNA Blood Mini kit. The complete HBV genome was amplified in 4 overlapping fragments (preS, S, X, and C) and phylogenetic analysis and distance calculations were performed by using MEGA v.4. Genotyping was performed by analyzing the complete genome or at least 1 of the 3 fragments of preS, S, or C genes. Subgenotyping was done on the full-length genome or on at least 2 complete fragments preS, S, or C. Sequences were submitted to EMBL/GenBank/DDBJ under accession nos. FJ692502–FJ692553 (Haiti S-fragment sequences), FJ692557–FJ692613 (Haiti complete genome sequences), and FJ692554–FJ692556 (Nigeria complete A5 sequences).

2 Results

2.1 Genotypes and subgenotypes

In 7,147 blood samples of pregnant Haitian women, HBsAg prevalence was 5%, ranging from 1.0% to 8.5%, depending on the sampling clinic. Of 320 HBsAg–positive samples available, 247 (77.2%) were positive for at least 1 of the 4 overlapping PCR fragments (Table 12). Interpretable sequences from at least 1 of the 4 PCR fragments were obtained from 213 viruses. A total of 179 of these strains could be clearly assigned to a genotype by analyzing the complete genome or at least 1 of the 3 fragments of preS, S, or C genes. Of the 213 strains, 31 showed signs of mixed infection or recombination, and 3 strains were considered outliers because they could not be genotyped.

Phylogenetic analysis of the above 179 genotypeable strains (excluding mixed, recombinant, and untypeable strains) showed that 128 (71.5%) viruses belonged to

genotype A; 40 (22.4%), to genotype D; and 11 (6.1%), to genotype E (Table 13). Genotype A strains were attributable to subgenotypes A1 (n=77 [43.0%]) and A5 (n=35 [19.6%]). Genotype D strains belonged to D4 (16.2%) and D3 (3.9%). Fifteen viruses of genotype A and 4 of genotype D could not be further subgenotyped (Table 13) because only partial gene sequences or single preS, S, or C fragments were obtained. In all of the above strains, genotypes of the different fragments agreed with each other. In addition, 31 viruses were suspected mixed genotype infections or recombinants; they were not included in the above analysis and are discussed later.

Table 12: Number of serum samples investigated, including suspected mixed and recombinant strains of HBV, Haiti

Sequence or gene	No. samples
No. sera (HBsAg pos) available	320
No. sera PCR pos	247
No. of sera of which sequences were obtained*	182
Full length genome	68
Full preS fragment	57 (37)
Full S fragment	67 (19)
Full X fragment	67 (20)
Full C fragment	20 (34)
No. sera of suspected mixed strains	25
Full length genome	10
Full preS fragment	8 (3)
Full S fragment	12 (2)
Full X fragment	10 (2)
Full C fragment	3 (3)
No. sera of suspected recombinant strains	6
preS fragment	3
S fragment	-
X fragment	3
C fragment	-

*No. serum samples includes 3 outliers and excludes mixed or recombinant strains.

Table 13: Prevalence of HBV genotypes and subgenotypes, excluding mixed or recombinant strains and untypeables, Haiti

Genotype or subgenotype	No. (complete genomes; partial strains) of genotypeable or subgenotypeable strains	Genotypeable or subgenotypeable strains, %
A	128 (63; 65)	71.51
A1	77 (36; 41)	43.02
A2	1 (1; 0)	0.56
A5	35 (21; 14)	19.55
D	40 (5; 35)	22.35
D3	7 (2; 5)	3.91
D4	29 (3; 26)	16.2
E	11 (1; 10)	6.15
Total	179 (69; 110)	100

2.2 Phylogenetic analysis

2.2.1 Subgenotype A1

Phylogenetic analysis of A1 complete genome sequences showed that Haiti strains form several clusters (not necessarily supported by bootstrap values) within available full-length A1 strains (Figure 14) from South Africa and other eastern African countries, as well as from the Philippines. Haiti's complete genome A1 strains showed a mean genetic diversity of 1.45% (maximum diversity of 3.86%) that rose to a mean genetic diversity of 2.49% (maximum 6.61% between FJ692589 and U87742) when all available A1 strains (mean 2.87%, maximum 7.64% between AY161140 and U87742) were included.

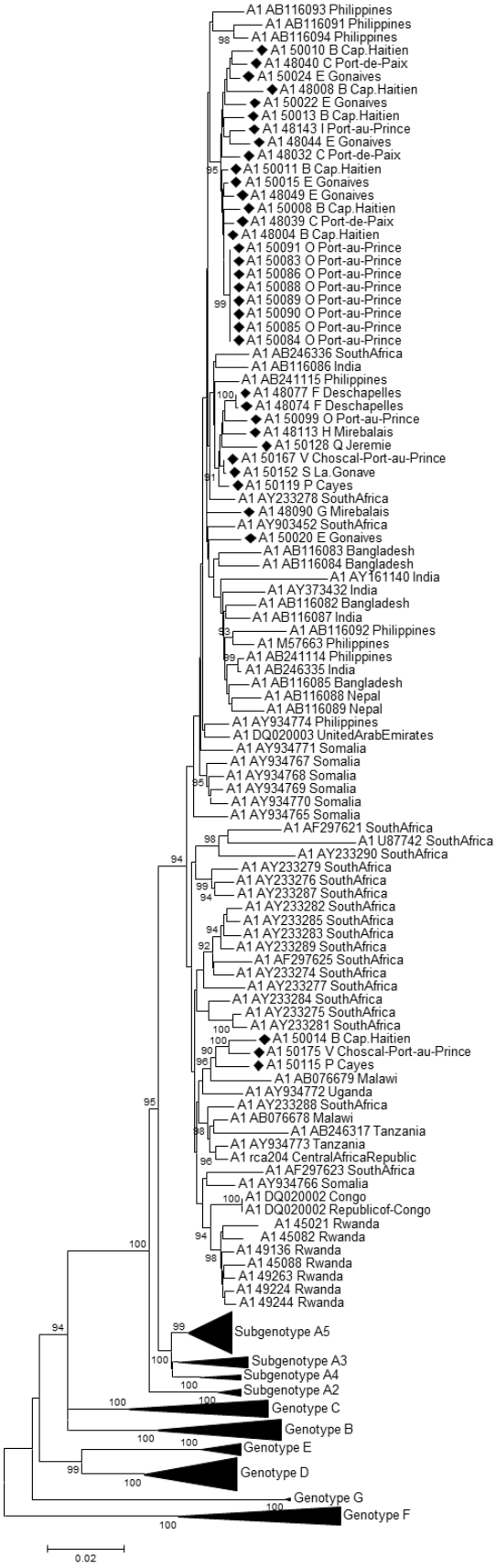

Figure 14: Phylogenetic analysis of selected sequences clustering with subgenotype A1, based on the complete genome.
Diamonds indicate Haiti sequences. All complete A1 sequences available in GenBank are included.

2.2.2 Subgenotype A5

A5 originally was proposed on the basis of the preS and preC/C gene fragments of 3 Nigerian strains (221) We present here the full-length sequences of the latter strains (accession nos. FJ692554–FJ692556) and compare them with all full-length A sequences from Haiti. Twenty-one sequences clustered with the only available A5 sequences from Nigeria. The overall mean intrasubgenotype diversity of A5 is 1.42% (maximum genetic diversity 2.89%). The mean intersubgenotype distance of the proposed A5 subgenotype was above the approximately 4% proposed for a new subgenotype (145, 217) for subgenotypes A1 (4.1%), A2 (4.8%), and A3 (5.1%); it was 3.8% when compared with the previously proposed subgenotype A4 (221). Nevertheless, A5 strains from Haiti and Nigeria form 2 distinct phylogenetic subgroups within A5, supported by high bootstrap values (99%; Figure 15). These subgroups are separated by a mean, minimal, and maximum genetic distance of 2.28%, 1.71%, and 2.89%, respectively.

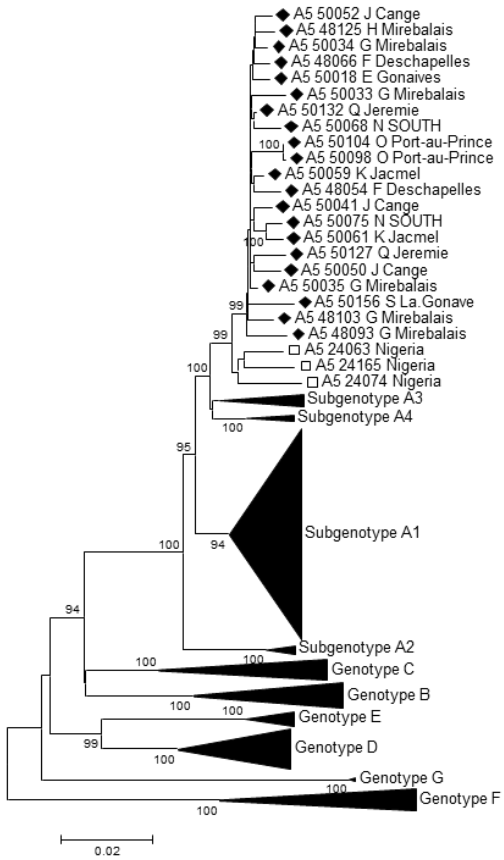

Figure 15: Phylogenetic analysis of selected sequences clustering with subgenotype A5, based on the complete genome.
Diamonds indicate Haiti sequences; squares indicate Nigeria A5 strains. All complete A5 sequences available in GenBank are included. Scale bar indicates nucleotide substitutions per site.

2.2.3 Genotype D

Because of low numbers of complete genome sequences available for genotype D in Haiti, we analyzed this genotype on the S fragment. In Haiti, most D strains belonged to D4 (29/179 [16.2%]). Besides a small cluster of D4 sequences from Australia and Papua New Guinea, a few D4 sequences from Rwanda (accession nos. FM200194, FM200212, and FM200213) (117) and single D4 sequences from Spain and France are available in GenBank. D4 sequences from Haiti were closely related to those from the latter 3 countries and somewhat separate from the Australian cluster (Figure 16). Despite the relatively high prevalence of D4 throughout most of Haiti, the mean genetic diversity of D4 S fragment sequences was only 0.39% (maximum diversity of 2.18%).

D3 (7/179, 3.9%) was less frequently found in Haiti than D4. Sequences seemed to form small geographic and genetic clusters, 1 of which most closely resembled strains from Rwanda (GenBank accession nos. FM200190, FM200191, FM200197, and FM200205)(117) but all D3 strains in Haiti were interspersed among strains from Brazil (Figure 16). Haiti D3 S fragment sequences showed a mean genetic diversity of 0.74% (maximum diversity of 1.44%).

2.2.4 Genotype E

The prevalence of genotype E sequences was surprisingly low, with only 11 (6.1%) of the 179 subtypeable strains being classified as this genotype. Available S-fragment sequences clustered with African HBV/E strains and were interspersed as individual strains among those HBV/E strains (Figure 17). The mean and maximum genetic diversity of the Haiti S fragments were 0.76% and 1.93%, compared with 0.74% and 4.66% of all African genotype E strains.

2.2.5 Mixed Infections and Recombinations

We suspected mixed infections in 25 samples either because of at least 5 divergent nucleotides corresponding to discrepant (sub)genotypes in at least 1 overlapping region of the PCR fragments or because of divergent nucleotides within the fragments after additional PCR analyses. Mixed infections included all genotypes in Haiti, as well as 1 B4 (S-fragment) and 1 C (X-fragment) sequence. Six other strains showed possible recombinations within the preS fragment or the X fragment. PreS-fragment recombinants were based on HBV/E and HBV/A, whereas those in the X fragment emanated from HBV/G in recombination with genotypes D or A. One of

these strains also showed signs of mixed infection. Although sequences were relatively short, recombination breakpoints seemed to be located around nucleotide 800 on the X gene; the location varied in the preS gene (nucleotide 330, 640, or 870).

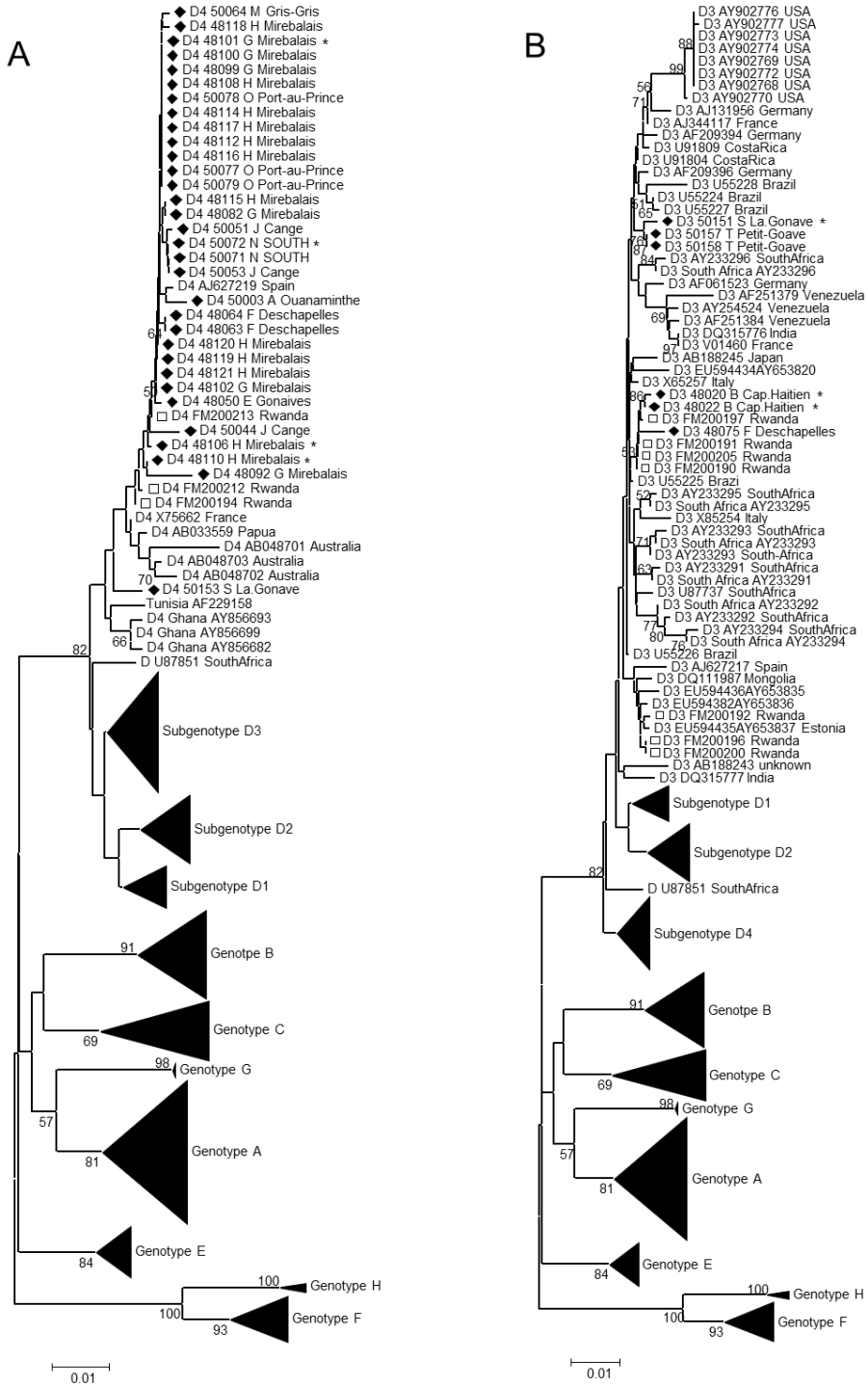

Figure 16: Phylogenetic analysis of selected sequences clustering with subgenotype D4 (A) or D3 (B), based on the S fragment.
Diamonds indicate Haiti sequences; squares squares indicate Rwanda strains. Asterisks indicate potential mixed or recombinant strains. Scale bar indicates nucleotide substitutions per site.

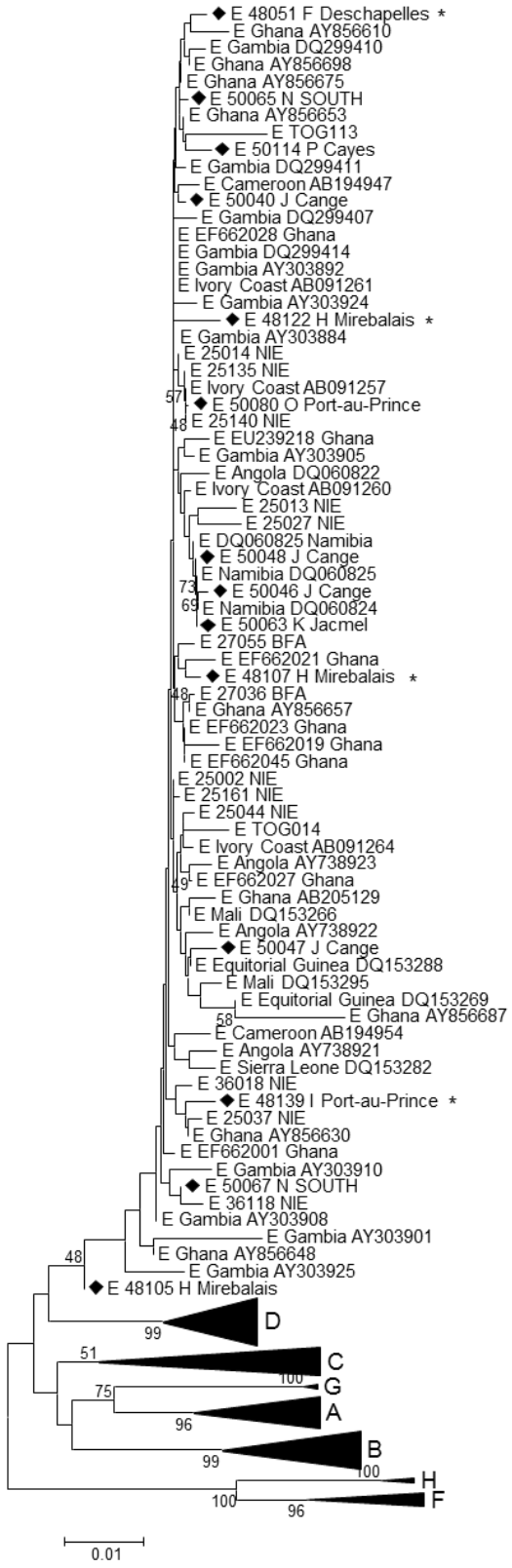

Figure 17: Phylogenetic analysis of selected sequences clustering with genotype E, based on the S fragment.
Diamonds indicate Haiti sequences. Asterisks indicate potential mixed or recombinant strains.

3 Discussion

More than 90% of today's Haitian population is descended directly from African slaves (64) exported from the late 17th century through the early 19th century (194). Because vertical transmission and household transmission during early childhood are important routes of infection and are associated with excess risk for chronic disease, HBV is transmitted between generations (48, 306). Thus, HBV strains in Haiti may to some extent reflect strains that were prevalent in Africa several centuries ago.

3.1 Subgenotype A1

Forty-three percent of African HBV sequences belong to subgenotype A1, the main African A subgenotype. This subgenotype was found in most eastern African countries (147) and dominated in this region for which larger sets of HBV strains have been characterized, including Somalia, Kenya (147), Rwanda (117) and South Africa (147). Subgenotype A1 is essentially absent from West Africa (i.e., west of the African E/A1 genotype divide) (117) and from other continents. During the peak of Haiti's slave importation during the late 18th century, almost 60% of captives came from southeastern and central Africa (92, 93) (Figure 18). Complete Haitian A1 strains (36/77 A1 strains) formed several small clusters within A1, at least the largest of which is supported by a bootstrap value of 95%. Thus, the phylogeny is highly suggestive of multiple early introductions into Haiti of distinct A1 strains from eastern Africa that continued to spread in Haiti's population.

3.2 A5, a new subgenotype

One third of HBV/A strains clustered with a group of rare strains that have been found only in southwestern Nigeria (221). Because only preS and preC/C sequences had been available, these strains had only provisionally been assigned to a new subgenotype, A5, until full-length sequences would become available. Complete A5 genome sequences from Haiti (n=21) and Nigeria (n=3) presented in this study showed a mean intrasubgenotype diversity of 1.42%. The mean intersubgenotype distances of A5 are above the approximately 4% of the definition of a subgenotype (3) for subgenotypes A1 (4.1%), A2 (4.8%), and A3 (5.1%). Compared with the proposed subgenotype A4 (3.8%), the mean intersubgenotype diversity is only slightly <4%. Thus, together with high bootstrap support (99%), these strains fulfil the formal definition, proposed by Kramvis et al. (147), of a new subgenotype A5.

A5 has been found only in southwestern Nigeria, the former Bight of Benin, where the first wave of slaves brought to Haiti originated (92). Interestingly, all A5 sequences from Haiti clustered together, but somewhat separately (bootstrap support of 99%) from those from Nigeria, suggesting early evolutionary separation of the 2 clusters. This distinct clustering further corroborates early introduction of these strains to Haiti during the slave trade. A5 strains in Haiti were considerably more homogenous than A1 strains, possibly reflecting their geographic confinement to and homogeneity in Africa.

Compared with subgenotype A5, A1 showed only a slightly higher mean genetic distance between African and Haitian strains (2.62% for A1; 2.28% for A5). However, when each Haitian A1 group is considered as a separate and independent introduction, the mean genetic diversity of these groups is only 0.8%–1.6% (for the different groups), indicating more recent introduction of these strains into Haiti.

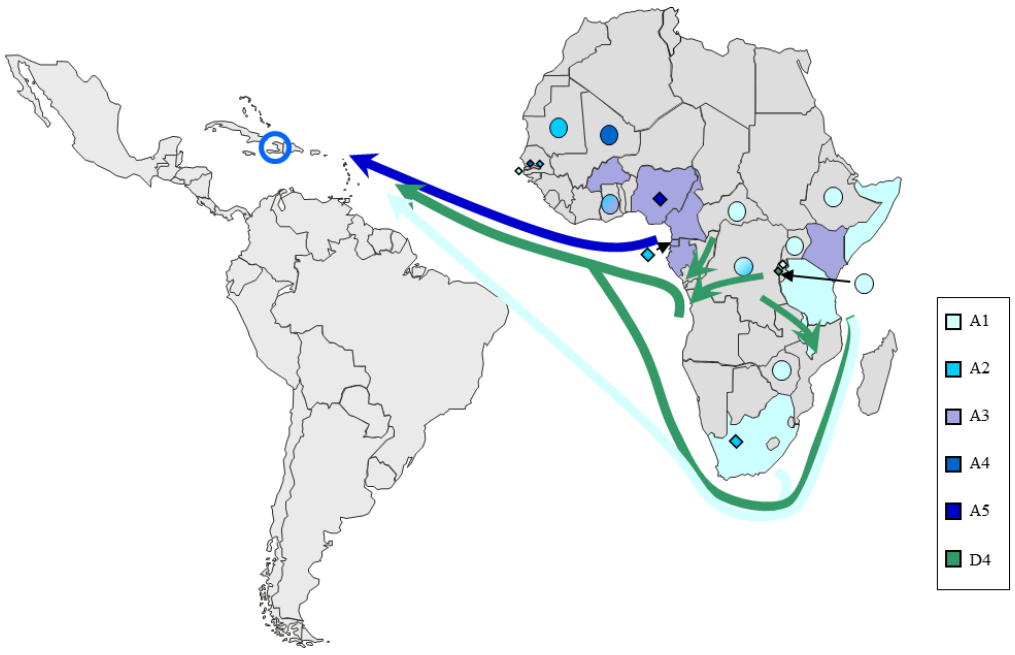

Figure 18: Distribution of hepatitis B virus A subgenotypes and D4 (only in Rwanda) in Africa and their potential routes of spread toward Haiti.
Colored dots indicate African countries with <10 A strains available; full color indicates countries with >90% dominance of 1 subgenotype; or a 60%–90% predominance of 1 subgenotype, with minority subgenotypes shown as diamonds. Subgenotypes other than A1 and D4 are not shown for Rwanda. Sequences included were obtained from GenBank and unpublished data.

3.3 Genotype E

We have extensively investigated HBV genotypes in Nigeria (197, 221) (and unpublished data). Of almost 300 sequences from both southern and northern Nigeria, 95% belong to genotype E and 4% to subgenotype A3. Thus, in Nigeria, the only location where HBV/A5 has been found, almost all HBV carriers are infected by genotype E, and the most prevalent none-E strains belong to subgenotype A3. In contrast, genotype A5 is rare and confined to Nigeria. Historical records (92) and the prevalence of A5 confirm that ancestors of the Haitian population came from the Bight of Benin, one of the most important slave trading posts. Nevertheless, only ≈6% of strains in Haiti belonged to genotype E. Single sequences were interspersed among current HBV/E sequences from Africa, with little genetic distance between them, suggesting that these HBV/E strains were introduced only recently into Haiti. The recent establishment HBV/E into Haiti strongly indicates that genotype E was essentially absent from West Africa when and where slaves were assembled for transport. Recent introduction of genotype E into the general West African population, as we have suggested previously (197, 221), also would explain the low genetic diversity of this genotype.

3.4 Subgenotype A3

A3, the minority subgenotype in West Africa, was virtually absent from Haiti, suggesting that this subgenotype also arrived later in the Bight of Benin. Interestingly, Cameroon is the only country where genotypes E and A3 cocirculate at similarly high prevalences (152, 197), suggesting that both E and A3 may cooriginate from this region.

3.5 Genotype D

More than 20% of sequences analyzed from Haiti belonged to genotype D (D3, D4). Subgenotype D4 strains are rare in the world, but we found a surprising 17% prevalence in Haiti. Interestingly, some of the D3 and the D4 strains were closely related to recent strains from Rwanda (117). With prevalences of 15.6% of D3 and 6.7% of D4 (117) Rwanda is also the only country where sizeable percentages of these 2 subgenotypes were found, further corroborating an origin of these strains from that part of Africa. During the second half of the 18th century, slaves were, to a large extent, collected in west-central Africa and shipped either from the western

coast (194) or the eastern coast (92, 194) to the Caribbean. Thus, an African origin of the D strains seems likely (Figure 18).

3.6 Time of evolution

Our results seem to agree with the time frame of the transatlantic slave trade. According to the simplest evolutionary model with a mutation rate of 4.2×10^{-5} (83), separation between Haitian and African A5 strains, with a mean genetic distance of 2.28%, would have occurred ≈270 years ago. When each Haitian A1 group is considered as a separate and independent introduction, the mean genetic diversity of 0.8%–1.6% corresponds to at least 100–190 years of evolution of each of the A1 groups in Haiti. These estimates for A1 and A5 seem to agree with the historical records that slaves (putatively infected with A5) from West Africa were introduced to Haiti ≈270 years ago, i.e., during the early phase (from the 1730s on) of the slave trade (92) and slaves (putatively infected with A1) from eastern Africa were exported to Haiti around the turn of the 18th century (92, 93).

The apparent absence of old genotype E strains in Haiti indicates that it was rare in West Africa at the time of the slave trade and emerged in the general African population by the beginning of the 19th century, after the majority of the slave trade was suspended. Indeed, the mean genetic diversity (1.74%) would evolve over the complete genome in only ≈200 years, even from a single virus.

Given the recent introduction of genotype E, the excessively high prevalence of this genotype throughout the genotype E crescent is difficult to understand. If HBV antibody prevalence was as high at the time of the emergence of HBV/E as it is today, why did genotype E spread so much more efficiently than genotype A (subtype A3 or A5) in West Africa? No evidence exists to indicate that immunity to genotype A (e.g., through vaccines) does not protect against genotype E. The rare A/E recombinants do not suggest a long cocirculation of both genotypes at a high prevalence. A more likely scenario is that HBV was relatively rare in Africa until genotype E was massively spread by a new route of transmission. Much evidence points toward mass injection campaigns performed in the Belgian and French colonies at the turn of the 19th century. Treatment campaigns against yaws during the 1920s–1950s (77, 266) and chemoprophylactic campaigns against sleeping sickness (89) were widespread and entailed sometimes reusing only a few syringes to treat, for example, 90,000 persons (89). Later, the extensive use of injectable

antibiotics, vaccines, and other drugs (e.g., against syphilis) with unsafe needles further promoted HBV transmission (77, 266). In Egypt, widespread transmission of hepatitis C virus has been linked to unsafe mass injection campaigns against schistosomiasis until the 1980s (87). Because HBV is estimated to be 10 times more transmissible than hepatitis C virus (87, 266), injection with unsafe needles is a possible route of transmission of HBV (genotype E). In addition, early potentially contaminated vaccine preparations, as well as insect vectors, might be culpable in the spread of HBV infection.

The high prevalence of genotype A5 in Haiti strongly indicates that predecessors of the Haitian population came from the Bight of Benin. However, subgenotype A3 and genotype E, highly prevalent today in this part of Africa, are essentially absent in Haiti. This lack strongly indicates that HBV/E emerged only later in the general African population. The high prevalence of HBV/E in large parts of Africa further suggests that HBV hyperendemicity is a recent phenomenon and probably the result of the extensive use of unsafe needles.

Part 4: Bayesian inference of the evolution of HBV/E

This manuscript has been submitted for publication as:

Andernach IE, Hunewald OE, Muller CP. Bayesian inference of the evolution of HBV/E.

I.E. Andernach contributed significantly to the study design and computational analyses and was in charge of data analysis and writing of the manuscript.

Except for Cameroon, where genotype A is dominant, genotype E is highly endemic in most of sub-Saharan Africa. Despite its high prevalence and wide geographic spread throughout large parts of Africa, HBV/E has a surprisingly low mean genetic diversity of only 1.75% over the full-length genome, in contrast to 4% diversity for the African HBV/A strains (6). It has been suggested that it would take only about 200 years for the genetic diversity of HBV/E to develop (8, 197), while recent studies indicate an even earlier emergence of HBV/E (5, 85). To understand the evolution of the virus, a number of studies have been performed to assess the nucleotide substitution rate of HBV. However, because of differences in the computational approach as well as the dataset, large differences of nucleotide substitution rates were observed, ranging from ~10^{-4} to 10^{-5} substitutions per site and year (s/s/y) (83, 141, 216, 222, 223, 320). As most of these studies were based on repeated sampling in the same chronic carriers of the virus or on viral strains collected from mother child pairs, the observed substitution rates are largely indicative of short evolutionary rates and may not necessarily reflect general longer term evolutionary rates. A recent study estimating HBV mutation rates on a selected dataset used Bayesian Markov chain Monte Carlo (MCMC) analyses and revealed a relatively high substitution rate of 3.7×10^{-4} to 7.7×10^{-4} s/s/y on the full-length genome (327). A substitution rate of 3.2×10^{-4} was furthermore reported for a small HBV/E dataset (5). All of these studies indicate an only recent introduction and a short time of evolution of HBV/E in the general population of sub-Saharan Africa.

Recently, a study investigated the long-term evolution of HBV worldwide on a selected HBV S-gene dataset, including all genotypes and subgenotypes (229). The authors linked the global spread of HBV to human migratory patterns to estimate a global evolution during the past 34,000 years and found a long-term substitution rate of $2.2x10^{-6}$ s/s/y on the S-gene which is almost 100 times slower than previous estimates. Based on this mutation rate, the tMRCA of genotype E was estimated to be 6000 years.

In our study, we address these apparent discrepancies and provide evidence that HBV/E has indeed only recently been introduced into the general West-African population. We performed phylogeographic analyses, including Bayesian MCMC modeling on extensive HBV/E full-length genome and S-gene datasets, in order to infer the timescale and dynamics of evolution of HBV/E. Our results confirm an only recent emergence of HBV/E, most likely within the last 130 years, and rapid expansion of HBV/E in sub-Saharan Africa most probably due to dramatic changes in routes of viral transmission.

1 Sequences

Non-recombinant sequences of HBV genotype E for which the country and year of sampling was available were obtained from NCBI (accessed December 2011). Recombinant strains were excluded using jpHMM and the RDP, GENECONV and MaxChi options of RDPv.3.44(182, 225, 268). The dataset included 167 full-length genome sequences and a total of 454 S-gene sequences with the accession numbers AB194947, AB194948, AB205188-AB205192, AM494689-AM494715, AM494717, AY738147, FN545821-FN545824, FN545827, FN545841, FN545842, GQ161755-GQ161766, GQ161768-GQ161774, GQ161776-GQ161787, GQ161789-GQ161805, GQ161807-GQ161845, GQ161847-GQ161812, GQ161814-GQ161836, HM363565-HM363611 (full-length) and AB194954, AB194955, AB205323-AB205329, AF323617-AF323619, AF323620, AF323621-AF323636, AJ604932-AJ605031, AM494719-AM494723, AM494725, AM494727, AM494730-AM494734, AM494737-AM494741, AM494743-AM494748, AM494751-AM494753, AM494832, FJ692540, FJ692544-FJ692548, FJ692550-FJ692553, FN547192-FN547199, FN547202-FN547205, FN547207, FN547209-FN547215, FN547217, FN547218, FN547220, FN547221, FN547225-FN547227, FN547229, FN547230, FN547232, FN547234, FN547236, FN547237, FN547239, FN547240, FN547248, FN547251-FN547253, FN547255-FN547257, FN547259, FN547261, FN547263-FN547272, FN547274-

FN547277, FN547279, FN547282-FN547288, FN547292-FN547294, FN547297-FN547299, FN547304, FN547310, FN547333, FN547339, FN547347, FN547356, FN547358, FN547359, FN547362, GQ161775, GQ161806, GQ161839, GQ161842, GQ161844, GQ161846, HM195104-HM195106, HM195108-HM195113, HQ385227, HQ385235, HQ385236, HQ385238, HQ385239, HQ385241, HQ385242, HQ385245, HQ385246, HQ385248-HQ385257, HQ385260-HQ385265, HQ385267 (S-gene only). Origins and sampling dates of the analyzed HBV sequences are summarized in Table 14.

The full-length and S-gene sequence sets were aligned using MAFFT v6 for Windows with the L-INS-i option.

2 Phylogenetic and phylogeographic analyses

Model estimation was performed for both the HBV/E full-length and S-gene datasets, using Topali v2 (192). Based on the Akaike information criteria obtained the general time-reversible model with four gamma categories and invariant sites (GTR+G+I) was selected for each dataset. In addition, the less complex symmetrical (SYM) and transversion (TVM) models were tested on the S-gene alignment. The analyses were carried out using the BEAST v.1.6.2 software package. Parameter settings were defined using BEAUTi v.1.6.2 with subsequent manual adjustments and analyses were performed with BEAST. Sequence sets were analyzed with a strict or relaxed molecular clock, using constant size, exponential and expansion growth tree priors. Additionally, geographical information was included as a "state" location parameter for a subset of parameterized runs. For each run all states before convergence and at least 10% of states, were discarded as burn-in. Each parameterized run was repeated to reach effective sample sizes (ESS) greater than 200. Runs were assessed using Tracer and merged using LogCombiner v.1.6.2. The most applicable parameterized merged runs were selected based on the highest Bayes Factors, as calculated by Tracer.

Tree calculation of the resulting output files was performed with TreeAnnotator v.1.6.2 and visualization with FigTree v1.3.1. Visualization of the geographic spread of HBV/E was performed on S-gene tree files including geographical information using SPREAD v1.0.3. Longitude and latitude data of the individual country midpoints were included in the analyses as geographical coordinates. Bayesian skyline analyses were performed in BEAST, using the Bayesian skyline coalescent of BEAUTi.

Table 14: Origin and sampling date of analyzed HBV genotype E full-length and S-gene sequences.

		No. analyzed sequences	
Country	Year	Full length	S-gene
Angola	2007	/	9
Benin	2001	/	13
Burkina Faso	2001	/	10
Cameroon	2007	2	11
	2006	1	1
	2005	/	3
	2002	/	19
	1994	2	4
Central African Republic	2004	28	55
Congo DRC	2001	/	3
	1998	1	1
Ghana	2000	5	15
Guinea	2006	77	79
Haiti	2006	/	10
Mali	2002	/	18
Nigeria	2007	47	47
	2006	4	40
	2005	/	33
	2001	/	15
	1998	/	8
	1997	/	9
	1996	/	3
Sudan	2009	/	26
Togo	2001	/	22
Total		167	454

3 Median-joining network (MJN)

A median-joining network of the 454 S-gene strains was constructed using Network v4.610 and visualized using Network Publisher. Briefly, MJN were constructed starting from minimum spanning trees that were combined within a single network. Aiming at parsimony, consensus sequences of 3 mutually close sequences were subsequently added at a time. These so called median vectors can biologically be interpreted as extinct ancestral sequences. This median operation then resulted in the most optimal network, reflecting the shortest possible distances between the individual nodes.

4 Results

4.1 Substitution rates and tMRCA

To estimate the time of evolution from a most recent common ancestor (tMRCA) and assess the nucleotide substitution rates, we performed Bayesian coalescent analyses on 167 HBV/E full-length and 454 S-gene strains. For both sequence sets the Bayes Factors (BF) significantly favored a Relaxed Molecular Clock model, with the Uncorrelated Exponential Clock having a higher BF than the Uncorrelated Lognormal Clock. Furthermore the Expansion Growth Coalescent with a UPGMA generated starting tree was favored.

Based on these parameters the substitution rate of the full-length HBV/E was estimated to be 1.86×10^{-4} substitutions per site and year (s/s/y), with a median tMRCA of 130 years (mean tMRCA of 174 years; 95% highest posterior density (95% HPD): 36-441 years). When including the geographic "state" parameter as a discrete phylogeographic inference, the substitution rate changed to 2.23×10^{-4} s/s/y and a median tMRCA of 116 years (mean 145 years; 95% HPD: 37-336 years), indicating an influence of the geographic "state" parameter in the analysis. While inclusion of this parameter is a prerequisite for visualization of the phylogeographic data, the resulting time estimations would have to be considered with caution. Since today's geographic boundaries were not present at the time of the supposed spread, one would expect the tMRCA without the geographic parameter to be more reliable with respect to time of evolution of HBV/E. Furthermore, calculations were performed using median values, as this would reduce the influence of rare outliers in the analyses.

Analyzing the S-gene sequences, the mutation rate of 1.9×10^{-4} s/s/y, as calculated for the full-length genome, would correspond to a median tMRCA of 71 years (mean 73 years; 95% HPD: 54-97) and 52 years (mean 53 years; 95% HPD: 39-71 years) when including the geographic "state" parameter. Based on a tMRCA of 130 years, as calculated for the full-length genome the mutation rate for the S-gene strains would correspond to about 7×10^{-5} s/s/y, (126 years tMRCA), while inclusion of the geographic parameter would increase the timeframe to about 156 years tMRCA.

4.2 Geographic distribution of HBV/E

Phylogeographic analyses of both the S-gene and full-length dataset showed that HBV/E strains formed several clusters. Interestingly, strains from individual countries did not necessarily cluster together, while strains sampled within one country at different time points could be found in the same, but also in different clusters (Figure 19, Figure 20). The analysis furthermore revealed a putative origin of HBV/E in the area of Nigeria. This is supported by spatial phylogenetic reconstruction, that revealed a clustering of HBV/E primarily in the region of Nigeria and a putative spread along the West African coast and to Angola, Congo DRC and the Sudan (Figure 21).

4.3 HBV/E population growth

The low genetic diversity and wide spread of HBV/E variants in sub-Saharan Africa suggests a rapid and recent increase in the number of HBV/E infections. Indeed, Bayesian skyline analyses of S-gene strains show a rapid population expansion over time. Based on a mutation rate of $7x10^{-5}$ s/s/y, corresponding to approximately 130 years of evolution from a MRCA, the most significant expansion of HBV/E infections would have occurred from the 1880s until the 1930s (Figure 22). This rapid expansion is furthermore supported by the star-like topology of the median-joining network (MJN) analysis (Figure 23). The central clustering of strains originating most prominently from Nigeria and Guinea, countries that are several thousand kilometers apart, indicate a recent and rapid spread of HBV/E. Haiti, in contrast, does not show a central clustering but rather monophyletic clades with a tMRCA of 62,4 and 55,9 years.

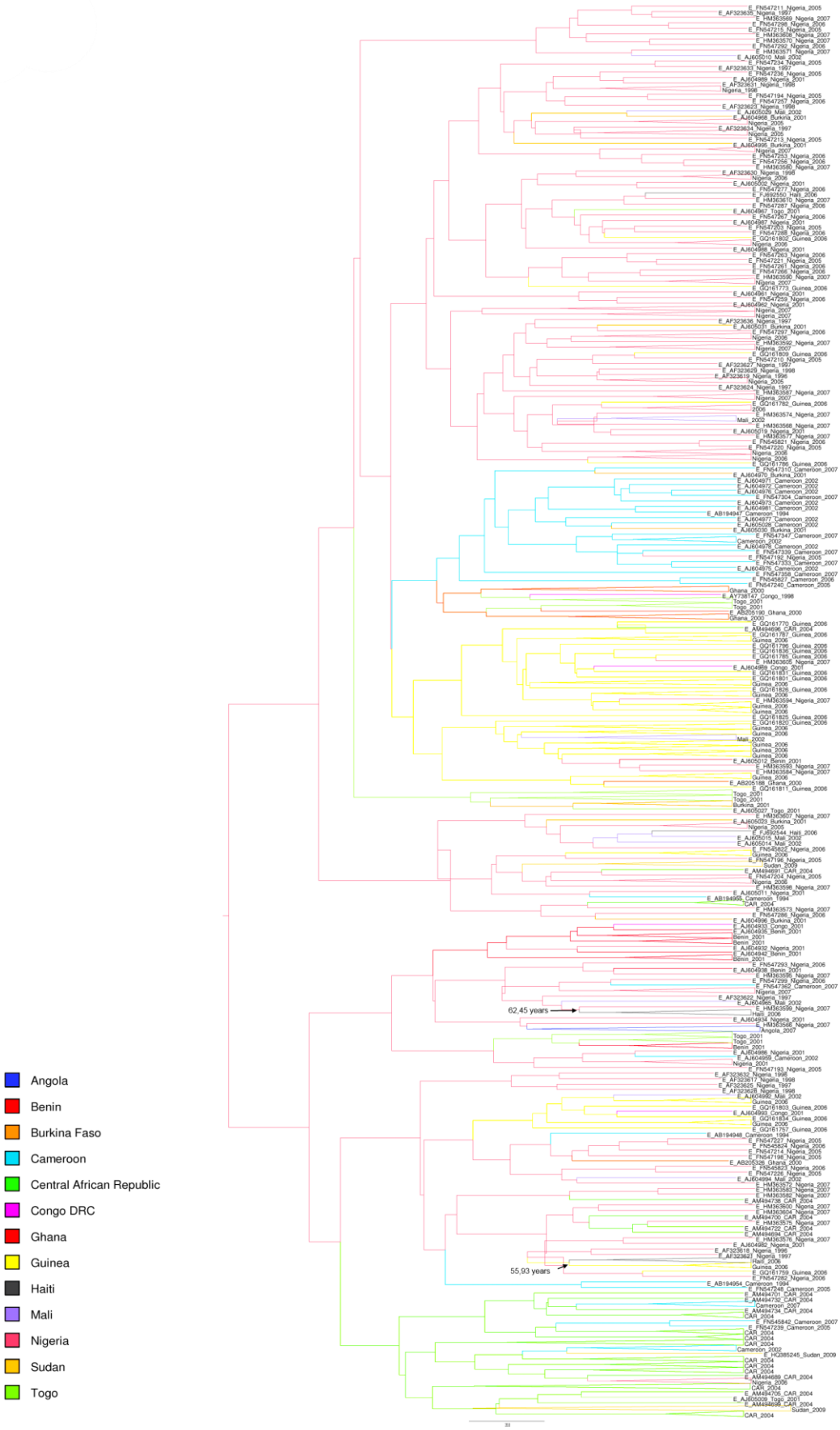

Figure 19: Phylogenetic analyses of all available HBV/E S-gene sequences, using the GTR+G+I model with geographic information.
Branching and roots of strains from individual countries are indicated by colors. Clusters with strains sampled in the same country and during the same year are collapsed.

Figure 20: Phylogenetic analyses of all available HBV/E full-length sequences, using the GTR+G+I model with geographic information.
Branching and roots of strains from individual countries are indicated by colors. Clusters with strains sampled in the same country and during the same year are collapsed.

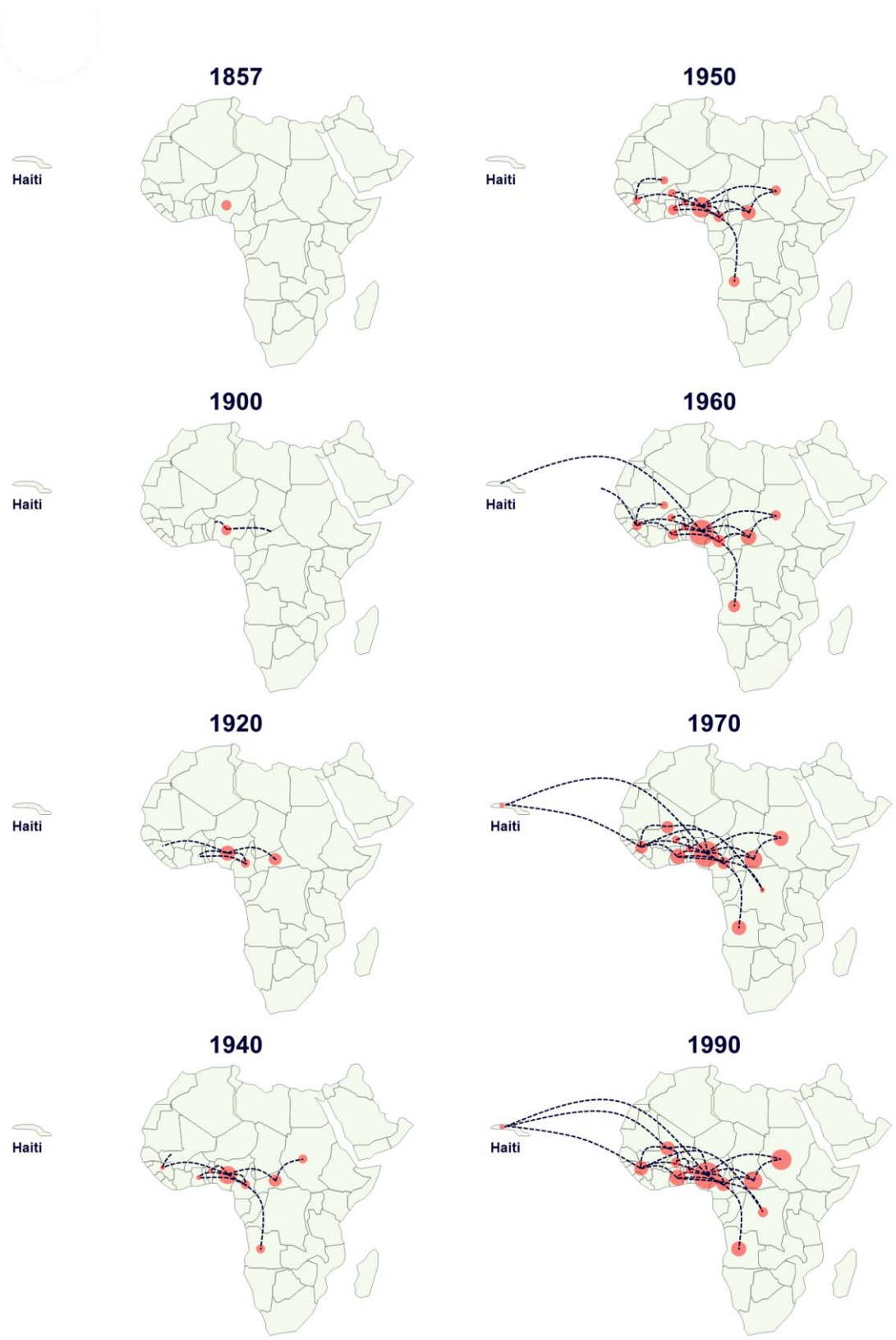

Figure 21: Schematic representation of the phylogeographic spread of HBV/E. Geographic and temporal spread of HBV/E strains for which the S-gene sequence is available, using a mutation rate of $7x10^{-5}$ s/s/y with the GTR+G+I model with geographic information.

5 Discussion

We performed extensive phylogenetic as well as phylogeographic analyses to characterize the origin and evolution of HBV/E. Such studies have obvious limitations due to sampling bias, since sampling is necessarily fragmented and incomplete. Nevertheless, under these assumptions, and including all available HBV/E strains, our analyses suggest a putative origin of HBV/E in what is today Nigeria, irrespective of whether the S-gene or the full-length dataset was analyzed. This finding is robust even when the oldest sequences from Nigeria are removed (data not shown). From this region the virus seems to have spread along the West African coast in the region from Guinea to the Central African Republic and subsequently towards Eastern and Southern countries, such as Sudan, Angola and the Democratic Republic of the Congo (Figure 19, Figure 20).

Median-joining network (MJN) analysis of this pool of genetically similar or even identical viral variants, which circulate all over sub-Saharan Africa, reveals a star-like topology, indicating a centralized origin of HBV/E (Figure 23). While the MJN analysis cannot give information on the temporal spread of HBV in Africa, the tight clustering of strains from distant countries several thousand kilometers apart, with sequences from Nigeria and Guinea dominating in the center of the MJN, further support a recent and rapid spread in countries along the West African coast. Indeed, Bayesian analyses revealed a median time of evolution from a most recent common ancestor (tMRCA) of 130 years, with a substitution rate of 1.9×10^{-4} substitutions per site and year (s/s/y).

In contrast to most other studies these values were based on an approach with limited priors to avoid unnecessary biases. In particular, we did not assume a specific time of evolution or mutation rate. This, together with the limited sampling period (13 years), would explain to a large extend the high confidence interval (36-441 years) of the 130 years tMRCA.

The uncertainty introduced by the high confidence interval is difficult to overcome by computational methods without the possibility of a deeper calibration with older viral strains. However, since HBV is transmitted by chronic carriers from generation to generation, our earlier HBV study from Haiti with its large population of descendants of African slaves provides additional independent clues. While the prevalence of HBV/E was conspicuously low in Haiti (6.1%) (8), the 10 S-genes of HBV/E strains were

interspersed within the African strains clustering with strains from Nigeria and Mali. They seemed to be the result of multiple introductions only after the HBV/E spread along the West African coast (Figure 19), with an observed tMRCA of the monophyletic Haitian clades of 62.4 and 55.9 years based on the above substitution rate of $1.9x10^{-4}$ s/s/y. These recent introductions are supported by the low variability in the different HBV/E cluster from Haiti in the phylogenetic tree. In contrast to strains from Africa that form regional clusters, strains from Haiti do not cluster together, but seem to originate from different African cluster and even from different countries, supporting the conclusion that individual HBV/E strains would have been introduced to Haiti only recently. This, as well as the only rare strains of HBV/E detected in the Americas (6), indicates that HBV/E was introduced and also spread in the general West-African population only recently and after the end of the transatlantic slave trade.

The only recent introduction of HBV/E into the general West-African population is furthermore supported by Bayesian skyline analyses that indicate an extensive increase in effective numbers of HBV/E infections (Figure 22) and is indicative of a recent and rapid spread of the virus explaining today's hyperendemicity in this region.

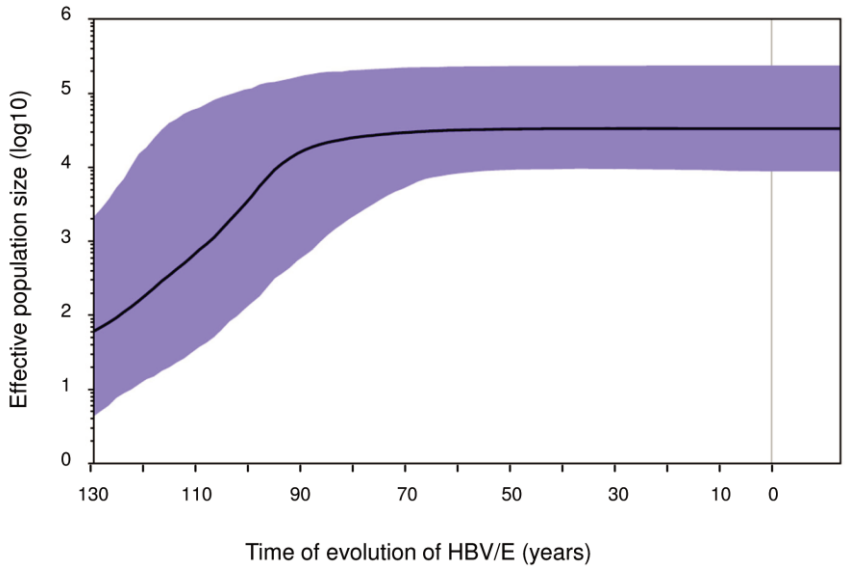

Figure 22: Bayesian skyline plot showing the epidemic history of the HBV/E S-gene dataset.
The plot indicates the median estimate of the effective population size, with the 95% highest posterior density indicated in blue. The applied timeframe ranges between the most recent sampling date and the calculated 130 years of evolution from the most recent common ancestor, as calculated in the HBV/E full-length analysis.

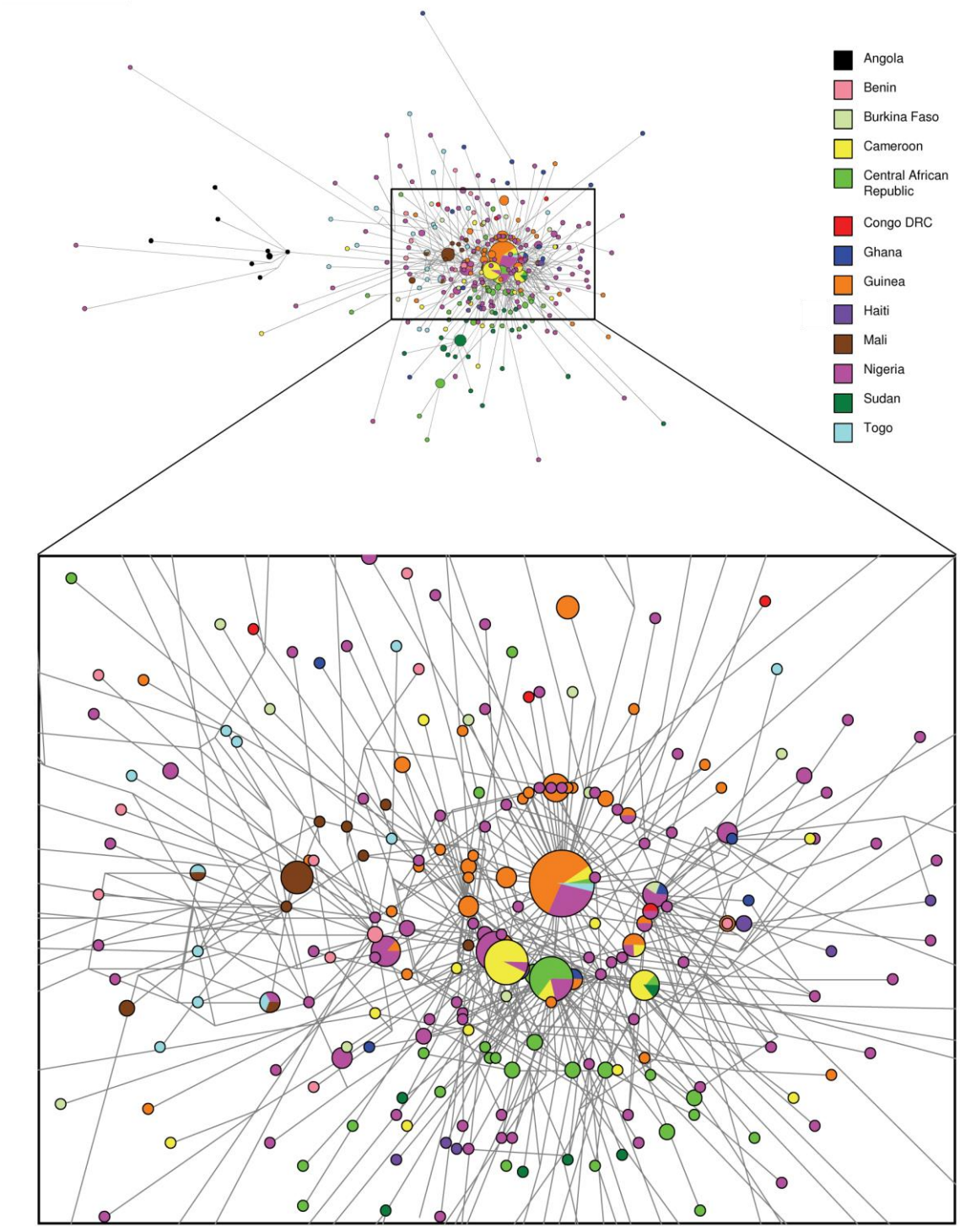

Figure 23: Median-joining Network of HBV/E S-gene sequences.
Pie charts represent sequence variants at the nodes, with colors indicating the country of sampling of individual sequences, the sizes reflecting the frequencies of the corresponding variants.

Previous studies estimated substitution rates of 7.72×10^{-4} s/s/y for HBV in general (327) and 3.2×10^{-4} to 4.3×10^{-4} s/s/y when analyzing individual HBV genotypes (5, 327). The latter substitution rates are comparable with those observed for HBV/E in our study. However, a recent study by Paraskevis et al. linked the evolution of HBV worldwide to human migratory patterns (229). The authors suggest that HBV co-expanded and co-migrated with human populations within the last 34,000 years and estimated a long-term substitution rate of 2.2×10^{-6} s/s/y for HBV in general and a tMRCA for HBV/E in Africa to be 6000 years (95% HPD: 3200 – 9400 years).

When applying their substitution rate to our HBV/E S-gene dataset, we found a tMRCA of 3717 years. This is, however, nowhere close to the 130 years of evolution observed in our study and is also incompatible with the absence of HBV/E in other parts of the world.

While methodological differences, such as the use of the S-gene or the full-length genome or differences in datasets may explain some of this discrepancy, biological considerations are equally important. Paraskevis et al. base their analyses on a limited dataset of selected S-genes of all HBV sub-/genotypes in order to infer a long-term substitution rate. In order to minimize mutational saturation, a conserved region may be appropriate for the study of long-term evolution, but not for HBV/E in Africa and for the estimation of (short-term) substitution rates. Because of the observed low genetic variability of HBV/E and because the S-gene does not fully represent the HBV variability (e.g. HBV genotypes and subgenotypes are defined on the basis of the full-length genome (147, 148)), our analyses on the full-length genome more closely reflect the time of evolution of HBV/E. Nevertheless, our results could at least be partially reconciled with those of Paraskevis et al. assuming that since its separation from the most closely related genotype D, HBV/E was confined for thousands of years to an isolated population from where it would have spread to the general West African population only 100 to 200 years ago.

In fact, based on a tMRCA of 130 years, the most prominent increase of HBV/E infections would have occurred approximately until the 1930s (Figure 22). As HBV is transmitted not only sexually, but also through percutaneous or parenteral contact with infected blood and body fluids (306), mass injection campaigns that became popular at the time provided a new route of transmission that may have disseminated the virus with unprecedented efficiency in the late 19th century and

well into the 20th century. Indeed, the above timeframe coincides largely with unsafe mass public health campaigns that were performed in sub-Saharan Africa at that time, such as smallpox vaccinations (111), treatment campaigns, e.g. against yaws and leishmaniasis (77, 266) and prophylactic injection campaigns against sleeping sickness (89). Even later the extensive use of injectable antibiotics, vaccines (e.g. against smallpox), and other drugs (e.g. against syphilis) with unsafe needles may have further promoted the transmission of HBV (77, 84, 232, 266).

The medical health campaigns, some even lasting into the second half of the 20th century, seem to be in contradiction with a saturation of the HBV/E population growth after the 1930s, as estimated in the skyline analysis. However, while an extensive spread in a (largely) HBV-naive region would lead to a rapid expansion of the overall viral population, the increasing prevalence of acute and chronic HBV carriers would subsequently result in a reduction of HBV-susceptible hosts which would, at least partially, explain a stabilization of the viral population despite continuing (local) transmissions.

Efficient horizontal transmission would accelerate viral replication and contribute to the high substitution rate observed. While the rapid spread of HBV/E can be explained by efficient horizontal transmission whether sexually, by unsafe mass injection campaigns or other practices, the current high contemporary prevalence of chronic carriers would be indicative of a high infection rate during early childhood.

While our study cannot provide a clear origin of HBV/E in Western Africa, Bayesian MCMC analyses corroborate a massive transmission and rapid expansion of HBV/E in the last 130 years. This is supported by MJN and skyline analyses, which showed a rapid increase of HBV/E infections, approximately until the 1930s. These results are in line with the low genetic diversity, the high prevalence of HBV/E in Africa and its absence in other parts of the world, indicating that the hyperendemicity of HBV in Western Africa is a recent phenomenon and likely due to dramatic changes in the routes of transmission in a relatively recent past.

Part 5: Characterization of Hepatitis Delta Virus in Sub-Saharan Africa

This manuscript is in preparation as:

Andernach IE*, Leiss L*, Tarnagda ZS, Tahita MC, Otegbayo JA, Forbi JC, Omilabu S, Komas NP, Gouandjika-Vasilache[f] I, Mbah OP, Muller CP. Characterization of Hepatitis Delta Virus in Sub-Saharan Africa

* These authors contributed equally to this work.

I.E. Andernach contributed significantly to the study design, experimental setup, data analysis, writing of the manuscript and was in charge of the training of the medical student.

Hepatitis D virus (HDV), a negative-strand RNA virus of 1.7kb size, is associated with hepatitis B virus (HBV) infection. Worldwide, about 15 million HBsAg carriers are estimated to be also infected with HDV (285). The co-infection with these two viruses results in fulminant hepatitis more frequently than HBV infection alone, whereas super-infection of HBV with HDV is associated with chronicity of HDV in up to 80% of carriers (310).

Of the eight distinct HDV clades (158, 241), clade 1 is the most prevalent clade worldwide and prevails in Africa, the Mediterranean basin, the Middle East as well as in Central and Northern Asia (73, 86, 118, 157). Clades 2 and 4 have so far been described in East and North-East Asia, with Clade 2 also present in the Yakutia region in Russia (121, 122, 312). Clade 3 has only been detected in South America (53). Clades 5 to 8 on the other hand have recently been described in West Africa and individuals that migrated to West Europe from this region (158, 178, 241). In this region HBV is highly endemic (6, 104, 147, 178, 197) and early childhood transmission is thought to be the most important route of infection (48). While this results in a high proportion of chronic HBV carriers, high rates of subsequent HDV super-infections add considerably to the burden of chronic liver disease (307).

Here, we analysed samples from Burkina Faso, the Central African Republic, Chad and Nigeria, where so far only rare and limited studies had been performed to assess HDV prevalence in HBV chronic carriers (159, 213). We analysed more than 2000 sera from these countries to determine the seroprevalence of HBV and HDV infection among different cohorts and analysed the genetic diversity and spread of HBV and HDV variants.

1 Clinical samples

Serum samples were obtained from apparently healthy individuals (Burkina Faso, Chad, Central African Republic), patients with symptomatic liver disease (Nigeria, Central African Republic) and HIV-positive individuals (Nigeria) between 1998 and 2010. These were stored at -80°C. The characteristics of the donors are shown in Table 15. All analyses have been approved by the competent ethical committees in each of the participating countries.

2 Serology

Hepatitis B surface antigen (HBsAg) was tested using either the Murex HBsAg Ver3 ELISA (samples from Burkina Faso, Nigeria, Central African Republic) or the Smart Check HBsAg Test) (Nigeria). Hepatitis D virus antibodies (HDV-Ab) status was determined using the Murex anti-Delta or the ETI-AB-DELTAK-2 ELISA. Equivocal samples were, if possible, retested and the result of the second test was accepted.

3 RNA and DNA isolation, RT-PCR, PCR and sequencing

RNA was extracted using the QIAamp Viral RNA Mini Kit. RNA was reverse transcribed into DNA and the complete genome was amplified in three overlapping fragments, covering nucleotides 307-870, 715-1302 and 868-483, respectively.

HDV quantification was performed, using 5µl of cDNA template material and the TaqMan PCR product of sample 36286 was cloned, using the TOPO-TA cloning kit. Concentrations were measured using the ND-1000 spectrophotometer and copy numbers were calculated. A 10fold dilution series of the plasmid, ranging from 10^7 to 10 copies, was used as a standard in the quantification reaction. In all reactions, correlation coefficient of the standard series was above 0.998.

HBV DNA was extracted from HDV TaqMan PCR positive sera of sufficient volume, using the QIAamp DNA Blood Mini Kit. HBV genotyping was performed on at least one of the HBV preS, S, X and C-gene regions. Quantitative real-time PCR was performed on 5µl of the extracted DNA and compared to a 10fold dilution series of WHO International standard 97/750.

4 Phylogenetic Analysis

Raw sequences were edited using SeqScape®, aligned with the L-INS-i option of MAFFT v6 and manually corrected. Neighbor-Joining trees and genetic distances were calculated using the MEGA 5 software. Statistical tests were performed using SigmaStat v3.11. All HDV full-length strains available on NCBI GenBank (accessed in June 2012) were included in the analysis.

Sequences were submitted to EMBL/GenBank/DDGJ under accession numbers JX888098-JX888135.

5 Results

5.1 Hepatitis B surface antigen serology

Of 2,273 leftover serum samples from apparently healthy adults (Burkina Faso, Tchad), children (Burkina Faso, Central African Republic), patients with symptomatic liver disease (Nigeria, Central African Republic) and HIV-positive individuals (Nigeria), 2,145 samples were tested for HBsAg, while for the remaining 128 there was not sufficient serum available for this test. 743 of the 2,145 were HBsAg positive, while 3 were equivocal and excluded from prevalence calculations.

The observed HBsAg seroprevalence rates ranged from 12.2% in children from Burkina Faso up to 81.7% in a liver patient cohort from Nigeria (Table 15). While the HBsAg prevalence was similar in healthy adults from Burkina Faso (14.5%, cohort size n=337 and 16.2%, n=370) and HIV-positive donors from Nigeria (14.6%, n=308 and 15.9%, n=69), it was different in children from Burkina Faso (12.2%, n=424) and the Central African Republic (44.3%, n=79). Interestingly, the latter was similar to the prevalence observed in liver patients from the same country (48.3%, n=29), while even higher prevalences were found in liver patients from Nigeria (62.9%, n=70 and 81.7%, n=126) (Table 15).

Table 15: Prevalences of hepatitis B surface antigen (HBsAg), hepatitis D antibodies (HDV-Ab) and hepatitis D RNA in cohorts from different West and Central African countries.

				HBV/HDV prevalence (no. positive/total. (%))		
Country	Sampling location (year)	Cohort	Cohort size (no.)	HBsAg	HDV-Ab*	HDV RNA**
Burkina Faso	Bobo-Dioulasso (2001)	Mothers	370	60/370 (16.2%)	1/40 (2.5%)	0/1 (0%)
	Bobo-Dioulasso (2001)	Children	424	52/424 (12.2%)	9/44 (20.5%)	0/9 (0%)
	Bobo-Dioulasso, Houndé (2007)	Pregnant women	337	49/337 (14.5%)	0/49 (0%)	-
Nigeria	Ibadan (1998)	HIV+	106	11/69 (15.9%)	3/11 (27.3%)	0/3 (0%)
	Lagos (2004)	HIV+	319	45/308 (14.6%)	3/45 (6.7%)	2/3 (66.7%)
	Ibadan (2003)	Liver patients	93	44/70 (62.9%)	3/44 (6.8%)	0/3 (0%)
	Ibadan (2006)	Liver patients	126	103/126 (81.7%)	1/78 (1.3%)	1/1 (100%)
	Abuja, Nasarawa state (2006)	HBsAg+	330	-	40/326 (12.3%)	15/40 (37.5%)
Chad	Military camp (2007)	Military personnel	50	14/50 (28%)***	-	3/14 (21.4%)***
Central African Republic	Bangui (2007)	Children	81	35/79 (44.3%)	1/35 (2.9%)	0/1 (0%)
	Bangui (2009)	Liver patients	37	14/29 (48.3%)	7/14 (50%)	5/7 (71.4%)

* Prevalence among HBsAg-positive samples, excluding equivocal samples.
** Prevalence among HBsAg-positive, HDV-Ab positive samples.
*** Prevalence calculations based on HBV and HDV TaqMan PCR. Samples positive for at least one assay were considered HBV positive.

One-hundred-twenty-eight sera were not tested with HBsAg ELISA but the burden of infection was assessed using the results from the HDV-Ab ELISA and the quantitative HBV and HDV PCRs. In particular, none of the 50 sera from Chadian military personnel were tested for HBsAg.

Among these 50 sera from the Chad, 13 were found to be HBV DNA positive while one HBV DNA negative sample was HDV RNA positive, resulting in an overall HBV prevalence of 28% (14/50) in the Chad (Table 15).

5.2 HDV-Ab serology

HDV-Ab ELISA was performed on 695 of the 743 HBsAg positive sera and 98 of the 128 sera that had not been tested for HBsAg. The remaining 48 and 30 sera could not be screened with HDV-Ab ELISA due to low sample volume. 68 of the 695 HBsAg positive and 3 of the 98 sera of unknown HBsAg status were HDV-Ab positive, while 9 and 1, respectively, were equivocal and excluded from prevalence calculations.

The HDV-Ab prevalences among HBsAg-positive sera from the different cohorts varied substantially (Table 15). They ranged from 0% in pregnant women (n=49) from Burkina Faso to 50% in liver patients (n=14) from the Central African Republic (Table 10). In Burkina Faso, the HDV-Ab prevalence in healthy adults (n=49 and 60) was generally low (0% and 2.5%), while a surprising 20.5% of children (n=44) were found to be positive. In the Central African Republic in contrast, the HDV-Ab prevalence was low in children (2.9%, n=35), while 50% of HBsAg-positive liver patients (n=14) (plus one of unknown HBsAg status) had antibodies against HDV (Table 10).

Also in Nigeria, HDV-Ab prevalences varied between the different cohorts (Table 15). In HIV-positive/HBsAg-positive donors from Lagos (n=45) and Ibadan (n=11) 6.7% and 27.3% were HDV-Ab positive, while in liver patients from Ibadan (n=78 and n=44) HDV-Ab were only found in 1.3% (plus one HBsAg-negative patient) and 6.8% of patients. A HDV-Ab prevalence of 12.3% was furthermore found in HBsAg positive sera from Abuja and Nasarawa state.

Of 50 sera from the Chad, none of which was analysed for HBsAg, 31 were screened for HDV-Ab, while the remaining 19 were directly subjected to RNA extraction. 2 of the 31 unselected samples (6.5%) were found to be HDV-Ab positive. All HDV-Ab positive samples and 30 untested samples underwent RNA extraction.

5.3 HDV quantification and genotyping

HDV quantification of 88 samples with complete HBsAg and HDV serology and 40 samples with incomplete or missing serology revealed 24 and 4 samples to be positive for HDV RNA. The viral load (detection limit: $1x10^3$ copies/ml serum) of the 28 HDV TaqMan positive samples ranged from $8.36×10^3$ to $1.63×10^9$ copies/ml without notable geographic differences in RNA titers.

22 of the HDV TaqMan positive samples, as well as 2 that were repeatedly TaqMan negative, were PCR positive for at least one HDV fragment. Fragment 307-870 was obtained for 19 samples (and partially for 3 samples), fragment 715-1302 for 21 samples and fragment 868-483 for 18 samples (plus 3 partial fragments). Taken together, the region 907-1265, commonly used for genotyping, was obtained for 22 samples and for 19 of these the complete L-HDAg ORF (nucleotide 954-1598) was available (and partially for 1 sample lacking 21nt). For a total of 16 samples, the full HDV genome was obtained. All 24 strains could be genotyped and revealed a predominance of clade 1 (17/24, 71%) and the presence of clades 5 (5/24, 21%) and 6 (2/24, 8%) in this study.

Clade 1 strains were found in sera from Nigeria (8/15, 53%), the Central African Republic (6/6, 100%) and Chad (3/3, 100%), while clades 5 (5/15, 33%) and 6 (2/15, 13 %) were exclusively found in Nigeria.

5.3.1 Phylogenetic analysis of HDV

Phylogenetic analyses of all available 87 HDV-1 full-length genome sequences from public databases (including the 12 strains from this study) revealed that HDV-1 formed multiple clusters in the phylogenetic tree (Figure 24). Interestingly, one cluster consisted of strains from all over the world, while 2 smaller clusters contained predominantly African strains, including our analysed strains (Figure 24). This apparent segregation of clade 1 could however not be clearly reproduced when analysing only the region nt 907-1265, typically used for HDV genotyping (Figure 25). Furthermore, the genetic variability of the "African" clusters (12.7% and 14% mean genetic diversity; 18.1% and 18.8% maximum genetic distance) was similar to the one observed among the other HDV-1 strains (12.8% mean; 21.3% maximum). The overall HDV-1 mean genetic diversity was 15.2% (23.4% maximum genetic distance).

Clade 5 was found in 33% of genotyped strains, but exclusively in samples from Nigeria. These formed a separate cluster apart from the other West African clade 5

strains (Figure 24 ,Figure 25). The overall mean intra-clade diversity of the full-length strains was 13% (16.9% maximum genetic distance). The intra-group diversity of the subgroups varied and was 3.9% (4.6% maximum genetic distance) in the Nigerian cluster and 12.8% (16.9% maximum genetic distance) in the remaining strains.

Also our two clade 6 strains from Nigeria clustered with a Nigerian (Figure 24 ,Figure 25). Analysing all HDV-6 sequences, we found a mean genetic diversity of 11.9% (14.6% maximum genetic distance).

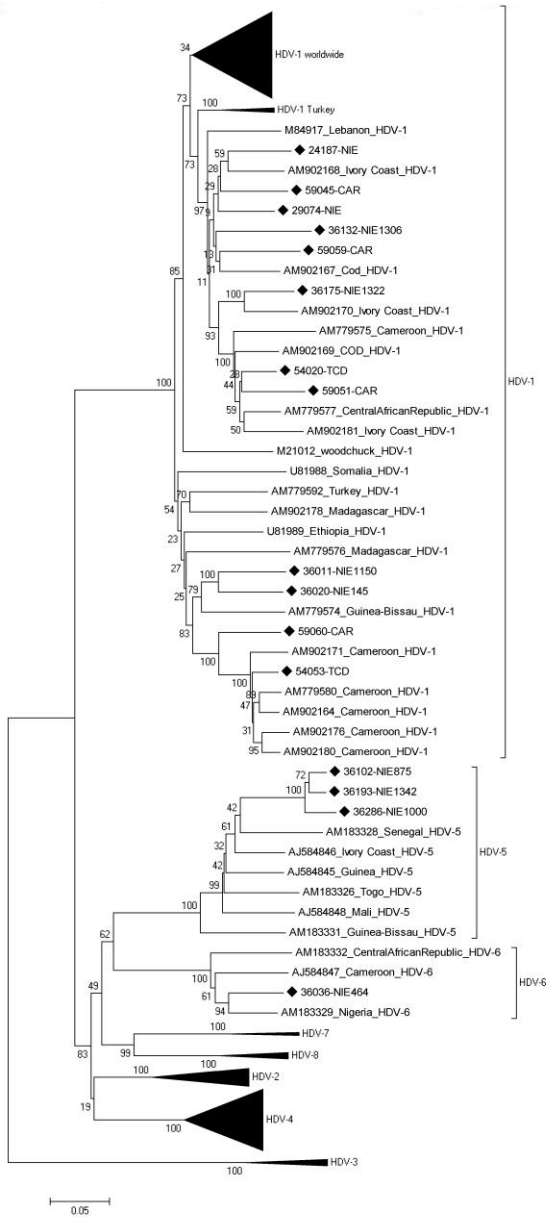

Figure 24: Phylogenetic clustering of all available HDV full-length strains.
Strains from this study are indicated by diamonds. Scale bar indicates nucleotide substitutions per site.

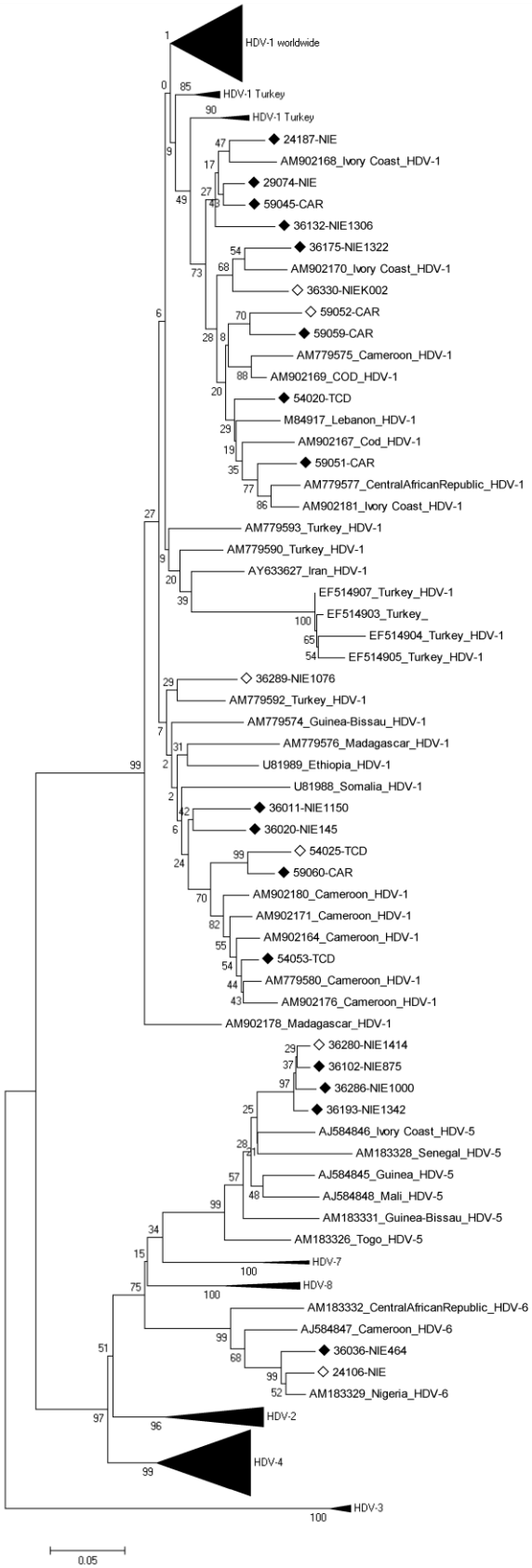

Figure 25: Phylogenetic analysis of the HDV genotyping region 907-1265 with reference sequences for which the full-length genome was available.
Diamonds indicate strains obtained in this study. Strains from this study for which the full-length sequence is available are indicated by solid diamonds. Scale bar indicates nucleotide substitutions per site.

5.4 Hepatitis B Virus quantification and genotyping

HBV quantification was performed in 23 samples with a known HDV genotype. In 11 of these (47.8%) HBV was undetectable, while in the remaining 12 samples $2.51x10^3$ to $5.32x10^9$ HBV copies/ml were detected. In 4 of these, the HBV load was even higher than that of HDV.

In 58.8% (10/17) of the HDV-1 positive sera, HBV was detectable, while in only 20% (1/5) of HDV-5 positive sera HBV DNA was found. HDV-1 positive sera seemed to have a higher HBV load (mean $4.79x10^8$ copies/ml; median $2.92x10^3$ copies/ml) than HDV-5 sera (mean $2.48x10^3$ copies/ml; median 0 copies/ml), although not statistically significant (p=0.311, Fisher's Exact Test). Furthermore, one HDV-6 sample was found to have $4.12x10^3$ HBV copies/ml.

Although HBV genotyping proved difficult due to the low viral load, seven HDV sequences were correlated to HBV genotypes. HBV-E and HDV-1 were found in 4 of the 7, while HBV-A/HDV-1, HBV-D/HDV-1 and HBV-E/HDV-6 were observed in one sample each.

5.5 Hepatitis Delta Antigen

Nineteen samples, for which the full HDAg ORF was obtained, were further characterized on the amino acid (AA) level. For 17 of these, the large HDAg was 214 AA long, while one sample each harboured an insertion (59045) or a deletion (36036) of one amino acid. For none of the strains a change in the reading frame was observed.

Clade 1 had been previously described to have higher virus assembly efficiency than other clades (160, 262). Indeed, all our clade-1 strains harboured proline 205, while in strains of the other clades arginine was found in this position. Similarly, the clathrin box, thought to be important for virion packaging (112, 113), varied between clade 1 (LFPSD) and our strains of the other clades (LPLLE). Furthermore, the isoprenylation signal, important for virion maturation (113), was CRPQ in clade 1 and CTPQ in our non-HDV-1 strains.

In addition AA 202 had been proposed as a marker to characterize the geographic origin of HDV-1 and was shown to largely vary between strains with African (serine 202) or Eurasian (arginine 202) origin (157). Indeed, our clade 1 strains revealed a serine at this position, except one strain from the Chad that had a proline 202.

6 Discussion

In the present study we characterized samples from Burkina Faso, the CAR and the Chad, from where no HDV data are available, as well as from Nigeria and characterized, for the first time, HDV strains from these countries. As only 34 HDV full length strains from Africa are publicly available, our 16 full-genome sequences considerably enlarge the genotype information from this continent.

In all our analysed cohorts HBsAg was highly endemic., but its prevalence varied widely, ranging from 12.2% in apparently healthy to 81.7% in symptomatic liver patients. Although it is difficult to compare the different cohorts, the HBsAg prevalence was similar in healthy adults and HIV carrier from Burkina Faso and Nigeria (14.5%-16.2%), while differences were substantial between children from Burkina Faso (12.2%) and the Central African Republic (44.3%). Although the cohort from the Central African Republic included children with a higher age (2-15 years) than in Burkina Faso (2-6 years), HBsAg prevalences in children from the Central African Republic were relatively constant and higher than 40%, irrespective of the age group analysed (2-5, 6-10 or >10 years of age), indicating a high burden and an increased early transmission of HBV in children from the Central African Republic. This high prevalence is however surprising, as HBsAg prevalences had previously been found to be 14% and 15.5% in young adults and high school students in the Central African Republic (144, 230). Furthermore, the observed HBV prevalence in the children was almost as high as in liver patients from the same country (48.3%), indicating an unusual route of transmission for the investigated group of children.

Also in military personnel from the Republic of Chad a high HBV prevalence of 28% was observed. As the HBV burden in the Chadian cohort was estimated based on HBV DNA and HDV RNA positivity, the HBV prevalence may even be underestimated in this group. Although generally military personnel seem to have higher HBV infection rates than unselected populations (181), HBV seems to be highly endemic in the Chad.

The rare recent studies that had been performed to assess the prevalence of HDV antibodies in the West African population, revealed varying HDV infection rates. While HDV-Ab prevalence in HBsAg carriers had been found to be below 20% in Cameroon and Mauretania (86, 181), it was found to reach up to about 70% in Gabon (178). Interestingly, also the HDV-Ab prevalences observed in our study varied

widely and ranged from 0% to 27.3% in asymptomatic carriers and 1.3% to 50% in liver patients. Surprisingly, 20.5% of HBsAg-positive children from Burkina Faso (12.2% HBsAg prevalence) were also HDV-Ab positive, a prevalence almost 10-times higher than the one observed in the mothers. These children are at high risk to develop severe fulminant or chronic hepatitis. Vast HDV-Ab differences were furthermore observed in liver patients from Nigeria and the Central African Republic. While in Nigeria a large proportion of liver patients were HBsAg positive, only 1.3% and 6.8% of these were HDV-Ab positive, with prevalences similar or lower than those observed in the other Nigerian cohorts (6.7%-27.3% HDV-Ab prevalence). On the other hand, 50% of liver patients from the Central African Republic were HDV-Ab positive, indicating that in the Central African Republic, in contrast to Nigeria, HDV super-infection might be the cause of chronic hepatitis with frequent severe liver conditions.

The HDV viral load varied widely in the analysed samples ranging from 10^3 to 10^9 copies/ml. As infection with HDV suppresses HBV replication, HBV was undetectable in almost 50% (11/23) of HDV RNA positive donors. This would however not necessarily be the result of a more efficient HDV replication, as these samples had an only slightly higher HDV load (mean $3.68x10^8$ copies/ml, median $3.18x10^7$ copies/ml) than the HBV DNA positive ones (12/23; mean $1.54x10^8$ copies/ml, median $1.06x10^7$ copies/ml). While on average a higher HDV load was observed in clade 1 strains (mean $2.5x10^8$ copies/ml, median $1.3x10^7$ copies/ml) than in clade 5 (mean $4.19x10^7$ copies/ml, median $8.98x10^6$ copies/ml) or clade 6 strains (mean and median $5.9x10^6$ copies/ml), the low numbers of clade 5 and 6 strains may not be representative enough. In fact, 71% (n=17/24) of the genotyped strains were assigned to clade 1, with strains originating from Nigeria, the Central African Republic and the Chad. Furthermore, 21% and 8% of our strains belonged to clade 5 (n=5/24) and clade 6 (n=2/24), respectively, all of these strains originating from Nigeria.

Phylogenetic analysis of HDV-1 strains revealed several clusters, although not necessarily with high bootstrap support, with strains from this study largely clustering together with other African strains (Figure 24). While previous studies indicated an ancestral origin of clade 1 strains in Africa (324), the genetic diversity of HDV-1 was similar in the African clusters (12.7% and 14% mean genetic diversity; 18.1% and 18.8% maximum genetic distance) as well as among the other (non-African) HDV-1 strains (12.8% mean; 21.3% maximum). A recent study identified an amino acid on the HDAg to distinguish between isolates of African (serine 202) and Eurasian (alanine

202) origin(159). Indeed, almost all our clade 1 strains revealed a serine 202 in the HDAg, supporting the hypothesis that at least the majority of African HDV-1 isolates might represent an ancient subgroup that originated from this continent. Together with the recently characterized HDV clades 5 to 8 (158, 241), HDV sequences are overall more divergent in Africa than anywhere else in the world (113), indicating that, similar to hepatitis B virus, HDV might have originated in Africa (241, 286). Nevertheless, the origin and emergence of HDV remains unclear. While it has been discussed that HDV might have co-originated with the most simple hepatitis B virus in ducks (DHBV), HDV has so far not been found to naturally occur in ducks or in mammals beside humans (112), indicating that HDV might have originated in ancestral humans, potentially from Africa.

All studies on HDV in sub-Saharan Africa revealed a predominance of HDV-1 in a region where HBV genotype E (and to a lower extent the African HBV/A subgenotypes 3 to 7) prevails. However, a close co-evolution of HDV-1 and HBV/E seems unlikely, since HDV-1 can be found worldwide whereas HBV/E is largely confined to Africa. A recent study in Turkey found that all HDV-1 variants were associated with HBV genotype D (157). More than 50% of these belonged to the putative African group (serine-202), while the remaining strains harboured alanine-202 and clustered in the Eurasian group (157). The high prevalence of HDV-1 might however be explained by a more efficient spread of this clade. Indeed, HDV-1 displays some unique nucleotide and amino acid features in the large HDAg when compared with the other HDV clades, potentially leading to more efficient viral reproduction (262). However, this is speculative and requires additional detailed studies.

In conclusion, our study revealed high prevalences of both HBV and HDV in multiple countries and cohorts from sub-Saharan Africa. Although the prevalences varied in the different cohorts, HDV superinfection was excessively high in cohorts with severe chronic hepatitis, while others were virtually free of HDV. The high variability between individual cohorts and countries indicates that HDV evolution is very complex. While especially clade 1 seems to originate from Africa, a co-evolution with the predominant HBV genotype E in this region is unlikely. Nevertheless, in a large proportion of cohorts analysed here, HDV represent a serious public healthcare problem, calling for an increased screening for HDV antibodies and extensive vaccination against HBV in this region.

Part 6: A High Variability of Mixed Infections and Recent Recombinations of Hepatitis B Virus in Laos

This manuscript has been published as:

Andernach IE, Jutavijittum P, Samountry B, Yousukh A, Thammavong T, Hübschen J, Muller CP. A High Variability of Mixed Infections and Recent Recombinations of Hepatitis B Virus in Laos. PloS One. 2012;7(2):e30245.

I.E. Andernach contributed significantly to the study design and was in charge of the experimental setup, experimental work, data analysis and writing of the manuscript.

In highly endemic countries, particularly in Asia where hepatitis B virus (HBV) genotypes B, C and I and their subgenotypes co-circulate, mixed infections (18, 114, 162, 168) and recombinations have been described (38, 41, 67, 195, 220, 265, 303, 314). For instance, in China almost 50% of infections are mixed (114) and mixed infections have also been reported from Thailand (124).

Recombination events do not seem to be negligible artefacts of HBV co-infections, as recombinant strains have become the dominant variant in certain regions. For example subgenotype Ba, a B/C recombinant, prevails in major parts of mainland Asia, and a recombinant between genotypes C and D has become the predominant variant in Tibet (67). Furthermore, recently a recombinant variant that circulates in Laos and Vietnam has been proposed as a new HBV genotype I (106, 220, 248, 294). Nevertheless, only few systematic studies of HBV recombinants have been published from regions with intensive co-circulation of several HBV subgenotypes (220, 274, 303).

In Laos, with 8.7% of chronic carriers (126), multiple HBV sub-/genotypes co-circulate (220) and were shown to belong to genotype C (55.4%), with subgenotypes C1 (93.1% of genotype C strains), C5 (6.4%) and C3 (0.5%), genotype B (42.2%), with subgenotypes B4 (78.5% of genotype B strains), B2 (11%), B5 (9.8%) and B3 (0.6%) and

the suggested genotype I (2.4%), with subgenotypes I1 and I2 (220). As a result of the high prevalence of chronic HBV carriers and the co-circulation of multiple genotypes and subgenotypes, Laos is prone to generate recombinant viruses. Here, we systematically analysed mixed infections and recombinations in 42 HBV strains from HBsAg positive first time blood donors. Our study revealed an unusual cauldron of mixed infections as a breeding ground of multiple and highly variable recombination events that so far have not led to new dominant strains.

1 Clinical samples

DNA was extracted using the QIAGEN DNA Blood Mini Kit, from sera of 40 HBsAg positive, but otherwise healthy, rejected first time blood donors from Vientiane City and Central provinces of Laos (collected in 2004/2005) and of 2 donors from the North of Laos (2006).

2 Amplification and cloning

The HBV S gene was amplified, using forward primer P2f and reverse primer 979. PCR products were cloned, using the TOPO TA cloning kit. Bacteria were plated onto Luria-Bertani-Broth agar plates, in the presence of kanamycin and X-Gal. White colonies were picked and inserts were amplified by colony-PCR using M13fw and M13rv primers.

3 Sequencing and phylogenetic analysis

The M13 PCR products, with 10 to 31 clones per donor, were purified and sequenced as described before (221) with above primers. Phylogenetic analysis and distance calculations were performed, using the MEGA v.4 with the neighbour-joining method of the Kimura 2-parameter model and a γ value of 0.6, with 1,000 bootstrap replicates. Sequences were submitted to EMBL/GenBank/DDBJ under accession numbers: HE652134 - HE652863.

4 Recombination analysis

Bootscan analyses were performed using SimPlot v.3.5.1 with a window size of 200 bp and a step size of 20 bp. Recombination breakpoints were confirmed by comparing each recombinant sequence with consensus sequences with a threshold of 50% for

inclusion in the consensus nucleotide sequence (consensus-50) (265) of subgenotypes B1-B8, C1-C7 and I1, I2. Of these, the subgenotype B6 consensus-50 harboured one ambiguous nucleotide. Additionally, consensus-50 sequences of genotypes B, C and I were included in the analysis, reconstructed from the associated consensus-50 sequences of the above subgenotypes to exclude a bias of the differently sized sequence sets available. The individual fragments of the potential recombinant strains were confirmed by phylogenetic analyses.

5 Results

5.1 Genotypes and subgenotypes

In 9157 blood donors in Vientiane city and central provinces from 2004 and 2005 the HBsAg prevalence was 8.8%. Of the 498 available samples of HBsAg positive blood donors, 453 were found to be PCR positive, 446 of them at least for the S gene. Sequence analyses of these donors revealed at least 5 ambiguous nucleotides within the amplified S gene of 40 donors, indicative of infections with several HBV variants. In addition, 2 donors from Luang Prabang in the North of Laos were suspected to be mixed infected and included in the analyses. The S gene PCR products from these 42 donors were further characterized by extensive cloning. Phylogenetic analyses revealed that among the 730 clones of these 42 donors, 16 donors had only clones of the same subgenotype (Table 16). Of these 16 non-mixed/non-recombined samples, strains from 12 donors belonged to subgenotypes B4 and one each to C1, and I1. Genotype B strains of the remaining 2 donors could not further be subgenotyped (Table 17; nomenclature according to Huy et al. (119)). The Asian recombinants, assigned to subgenotype B2-B4 (former Ba), and the proposed genotype I are not considered recombinants for the purpose of this study.

5.1.1 Mixed, non recombinant

The other 26 donors, or 5.8% of the 446 donors that were PCR positive for the S gene, were mixed infected with different variants. For 9 of these donors (14 to 30 clones) all clones were clearly assigned to a sub-/genotype, whereas the remaining 17 donors (3.8% of the 446 S gene PCR positive donors) had at least one (and up to 27 clones of the 10 to 31 clones generated) with signs of a recombination within the S fragment (Table 16). In addition, recombination events were detected in 2 strains from Laos

already published on NCBI (accession nos. FJ023979, FJ023832) which were included in the analyses.

In the above 9 donors with mixed, non-recombined variants, at least one clone of 7 of the donors was assigned to genotype C1, while the second sub-/genotype varied and belonged to subgenotypes B4 (n=1 donor), C5 (n=1 donor), I2 (n=1 donor), and non-subgenotypable genotype B clones (n=2 donors), while for one donor each B1 and B4, or C5 and B4 was found in addition to C1. For two donors no clone was assigned to genotype C and clones were attributed to genotypes B2 and B4 as well as B4 and non-subgenotypable B clones (Table 17).

5.1.2 Mixed, recombinant

In the other 17 donors with mixed-infections at least one clone showed signs of recombinations, and at least one clone showed no such evidence. 16 of these donors had non-recombined clones, that were assigned to genotype C, 15 to genotype B and 1 to genotype I, with subgenotypes C1 (n=16), C5 (n=1), B4 (n=14) and I1 (n=1) (Table 17). Strains of one donor were attributed to genotype C, but clustered separately from all HBV/C subgenotypes. Additionally, genotype B clones from 3 donors could not be subgenotyped, as the reference strains for genotype B subgenotypes B3, B5, B7 and B8 clustered interspersed with each other and the investigated strains (Figure 26).

5.1.3 Deletions and insertions

The S gene of the 730 clones was 681nt long, except for 7 clones, consisting of genotypes C1, B4 and B which contained single nt deletions (3 clones in donors with mixed infected, non-recombinant clones and 4 clones in donors with recombinant strains). For another 6 clones, consisting of genotypes B4 and B, the S gene sequence could not be fully recovered.

Table 16: Number of donors suspected to be mixed infected with at least 2 sub-/genotypes of HBV, Lao PDR.

Samples	No. donors
Detection-PCR positive	453
S-gene PCR positive	446
Total analyzed/suspected mixed infection	42
Non-mixed, non-recombinant	16
Total mixed	26
Mixed, non-recombinant	9
Mixed, recombinant	17

Table 17: Sub-/genotypes in HBV mixed infections in Lao PDR.

Category	Sub-/genotype(s)	No. donors
Donors suspected/analyzed	n.a.	42
Non-mixed, non-recombinant donors (n=16)	B	2
	B4	12
	C1	1
	I1	1
Clones in mixed, non-recombinant donors (n=9)	B4, B	1
	B4, B2	1
	B4, C1	1
	C1, B	2
	C1, C5	1
	C1, I2	1
	B1, B4, C1	1
	B4, C1, C5	1
Clones in mixed, recombinant donors (n=17)	B	1
	B4, B, C1	2
	C1, C5, B4	1
	B4, C1	11
	C1, C	1
	C1, I1	1

n.a.=not applicable

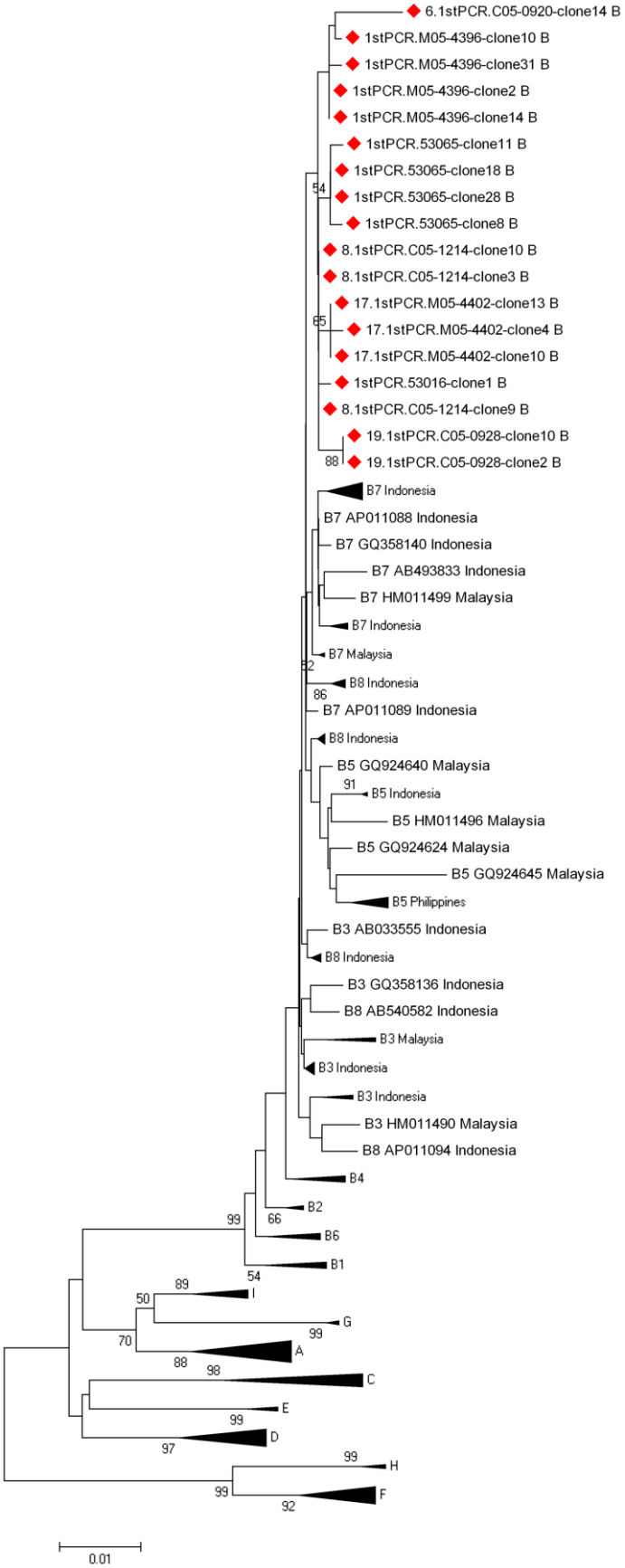

Figure 26: Phylogenetic clustering of selected HBV/B subgenotype B3, B5, B7 and B8 strains from GenBank and selected genotype B strains from Laos.
Non-subgenotypable clones are indicated by red diamonds. Scale bar indicates nucleotide substitutions per site.

5.2 Control experiments

Considering the large proportion of recombinant strains detected in this study, we performed extensive tests to exclude that there may be recombinants due to PCR or other artefacts. Therefore, M13 products of cloned, characterized HBV/B and HBV/C strains were mixed in various ratios (1:25, 1:5, 1:1, 5:1 and 25:1) and concentrations (1 ng and 100 ng total DNA) and amplified using different elongation times (1 min elongation as in the study protocol and a reduced elongation time of 20 sec). PCR products were cloned by TOPO TA, amplified by M13 PCR and sequenced. In the 1 min elongation PCR only rare recombinants (<2%, 2/107 clones) were found (Table 18). Only when the elongation time was reduced to 20 sec the rate of recombinations increased to an overall 16% (18/111 clones). At 1 ng final concentration and 20 sec elongation 6.4% of clones showed recombinations. This increased to 39.4% when using 100 ng template DNA *and* the short elongation time (Table 18). Thus recombinant artefacts seem to depend on high template concentrations *and* short elongation times.

After PCR amplification of our clinical samples the highest final concentration of PCR product before cloning was 154 ng/ul, corresponding to a template starting concentration considerably lower than the nanomolar concentrations of DNA analysed in the above control experiments. Although we cannot exclude that some recombinants may have developed only when the concentration of template increased during the later amplification cycles, these recombinants would be relatively rare and unlikely to be picked up during the cloning. This confirms that under our experimental conditions only very few recombinants (of the observed 9.9% of recombinant clones) would be the result of PCR artefacts. We also excluded that recombinations may have occurred during cloning. Thus, the observed high prevalence of recombinant clones in 65.4% of mixed infected donors is essentially free of PCR dependent or other artefacts.

Table 18: Number of PCR-induced recombinant clones in control experiments.

	Template concentration	
Elongation time	1 ng	100 ng
20 s	6.41% (5/78 clones)*	39.39 (13/33 clones)
60 s	2.53% (2/79 clones)	0% (0/28 clones)

*Clones for each condition originate from HBV/B and HBV/C strains mixed for amplification in ratios 1:25, 1:5, 1:1, 5:1 and 25:1.

5.3 Distance calculations of non-mixed, non-recombined samples

Distance calculations were performed in clearly subgenotypable clones of individual donors, with at least 5 clones of the respective subgenotype, including only clones for which the full S-gene could be recovered. These clones formed quasispecies within individual donors and revealed relatively low mean intra-group diversities for subgenotypes B2 (0.47%; n=1 donor) and B4 (0.08–0.47%; n=24), subgenotypes C1 (0.05–0.82%; n=11) and C5 (0.04%; n=1), as well as subgenotypes I1 (0.34%; n=1) and I2 (0.25%; n=1).

5.4 Amino acid analysis of clearly genotypable strains

The comparison of 651 clearly genotypable strains (excluding the 7 clones with nucleotide insertions that were not considered for analyses on the amino acid level) with consensus-50 sequences of genotypes B, C and I as well as their subgenotypes revealed multiple and variable amino acid (AA) changes over the S protein.

5.4.1 HBV/B clones

In 469 HBV/B clones, including the 6 clones of which the S gene sequence could not be fully recovered, AA substitutions occurred in 120 of 227 AA positions. In 7 of these positions substitution occurred in >5 clones: T27A/I/V (n=7 clones of d=5 donors), S61L/A (n=6, d=6), P62L (n=6, d=5), C76Y/R (n=8, d=6), P120S/T/L (n=20, d=14), M133I (n=31, d=2), T140I (n=8, d=6). Furthermore, a total of 6 internal stop codons (*) were detected.

5.4.2 HBV/C clones

In 154 genotype C clones we found AA substitutions in 56/227 positions, with 2 occurring in >5 clones: I92T (n=11, d=1), T118M/P (n=51, d=15). In the genotype C strains, a total of 4 internal stop codons were detected.

5.4.3 HBV/I clones

In the 28 genotype I clones, 19/227 AA positions were found to harbour substitutions. Due to a lower sample size of HBV/I strains, only 2 AA positions with >1 substitutions per site were further characterized. The investigated strains were found to harbour P111L/Q (n=2, d=2), I226N (n=12, d=1). In the genotype I strains, a total of 2 internal stop codons were detected.

5.4.4 Vaccine or treatment associated mutations

Despite the frequent amino acid substitutions only rare mutants had been described previously as vaccine or treatment associated mutations (172, 261, 315), including P120T (n=5, d=4), G145R (n=1), W172* (n=2, d=2), W182* (n=2, d=2), L192F (n=1), W196* (n=1), M198I (n=18, d=2) and W199* (n=1).

5.4.5 HBsAg subtypes

All 4 major HBsAg subtypes, that have been described previously in genotype B and C strains from Asia (209), were found among the analysed clones. In the 469 analysed genotype B strains HBsAg subtype ayw1 predominated (n=444 clones, d=35 donors), while 18 clones (d=3) were attributed to adw2 and one (d=1) to ayw2. Three clones (d=3) harboured 127S and were attributed to ayw. For three additional clones no HBsAg subtyping could be performed, as two clones harboured AA substitutions 122G or 160E and for one clone the HBsAg sequence could not be fully recovered.

The 154 genotype C strains belonged primarily to subtype adr (n=122, d=25), but also to adw2 (n=30, d=4), ayr (n=1, d=1) and ayw1 (n=1, d=1).

In the 28 genotype I strains on the other hand both adw2 (n=16, d=2) and ayw1 (n=12, d=1) prevailed.

5.5 Description of recombinant clones

A total of 72 recombinant clones were detected in the 42 analysed donors, from which 730 clones were generated. In 17 donors at least one recombinant clone was detected. All of the 72 recombinant clones exhibited one or two recombination site(s) within the S gene and revealed recombinations between genotypes B and C (Figure 27). Of the 185 genotypable fragments 72 were assigned to genotype B and 113 to genotype C.

In none of the analysed recombinant strains the recombination breakpoint could be exactly defined because homologous stretches of at least 17 nt length separated the recombined fragments (Figure 27). Interestingly, recombinations occurred all over the S-gene: nt 1–170 with 10 recombinant sites, nt 171–510 with 58 recombinant sites and nt 511–681 with 45 recombinant sites (Figure 27). For each of the individual 185 fragments of the 72 cloned recombinants the most similar strains were identified by phylogenetic analyses and pairwise comparison with a carefully selected dataset

that was composed of all available, clearly subgenotypable full-length genome strains from NCBI (accessed 08/2011) and all non-recombinant strains from this study.

Based on the simplest evolutionary model with a mutation rate of 4.2×10^{-5} mutations per site and year (83), the time of evolution from a hypothetical most recent common ancestor (tMRCA) was calculated. For 127 of the 185 fragments a non-recombinant strain of the same genotype was detected in the same donor. Two of the 127 fragments were most closely related (Kimura-2-parameter, pair wise genetic distances) to strains of the same donor (0% and 0.5%), with minimum genetic distances corresponding to 0 and 60 years of evolution from a hypothetical MRCA, while 58 fragments (of the 127) were genetically equidistant to clones found within the same donor and to strains from other origins (0–0.7%, 0–83 years to a MRCA). In 13 of the latter 2 and 58 fragments, the tMRCA was lower than or equal to the age of the donor. The remaining 67 of the above 127 fragments were most closely related to strains from another source (0–7.6%, 0–905 years to a MRCA).

58 of the above 185 fragments of recombinant strains, however, originated from donors that were *not* coinfected with a strain of the genotype found in the recombinant fragments. Of the latter 58 fragments 1 fragment was found to be most similar to a strain from Laos (0%, 0 years to a MRCA), and 56 fragments were equidistant to strains within and outside Laos (0–7.6%, 0–905 years to a MRCA). 1 fragment was most similar to strains outside of Laos (3.2%, 381 years to a MRCA).

In addition, recombination events were detected in the previously published strains FJ023979 and FJ023832 and the individual fragments (n=5) of these two recombinant strains were attributed to genotype B and C. For 1 of these fragments the most closely related strain was found within Laos (1%, 119 years to a MRCA), while 3 fragments were equidistant to strains within and outside Laos (0–7.6%, 0–905 years to a MRCA) and 1 fragment was most closely related to strains outside of Laos (0%, 0 years to a MRCA).

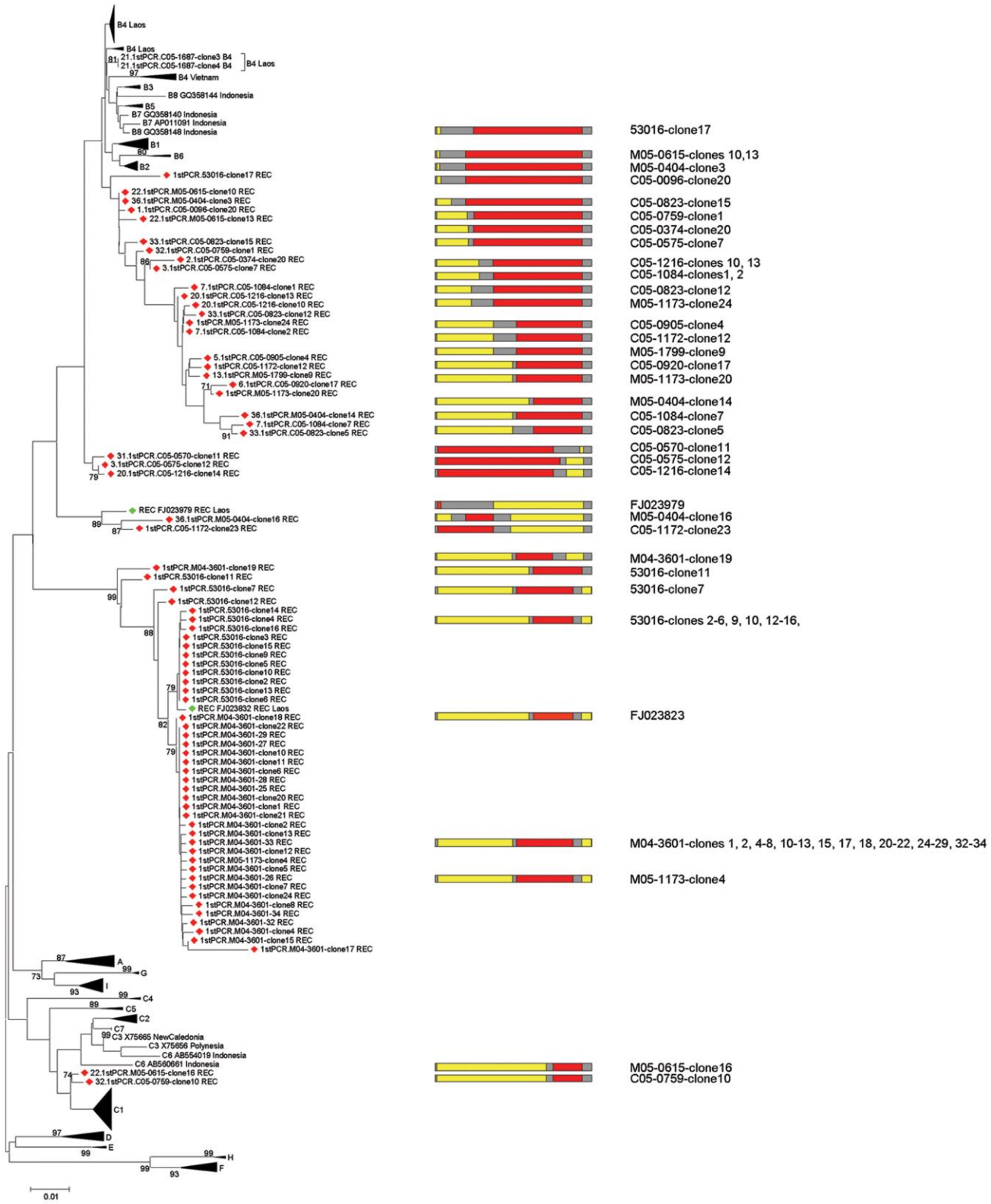

Figure 27: Phylogenetic clustering and recombination patterns of HBV recombinant strains.
Colour codes indicate genotypes B and C (red and yellow) involved in the recombination and regions homologous between at least 2 genotypes (grey). Recombinant strains from this study (red) or detected in previously published strains (green) are indicated by diamonds.

5.6 Amino acid sequence analysis of recombinant clones

Comparison of the 74 recombinant clones (including FJ023979, FJ023832) with the consensus-50 alignment revealed several consistent amino acid (AA) substitutions over the HBsAg. In these clones AA substitutions occurred in 47 of 227 AA positions. In 10 of these positions substitution occurred in >5 clones: P11H (n=44 clones of d=3 donors and FJ023832), T/A45P (n=43, d=3 and FJ023832), P/L49R (n=43, d=3 and FJ023832), C76F/S (n=43, d=3 and FJ023832), T118A/M (n=17, d=5 and FJ023832), F134I (n=42, d=3 and FJ023832), V190A (n=27, d=2), I218L (n=41, d=3 and FJ023832), F219S (n=41, d=3 and FJ023832), C221F (n=41, d=3 and FJ023832). Furthermore, 3 recombinant clones revealed internal stop codons. Of the observed amino acid substitutions only W182* had been described as vaccine or treatment induced mutation (172, 261, 315).

Similarly to the clear genotypable strains, the 4 major HBsAg subtypes were found within the recombinant clones. For 16 clones as well as FJ023832 and FJ023832 the end or the beginning of a recombinant fragment fell within the 'a'-determinant. For these clones the HBsAg subtypes were found to be ayw1 (n=14, incl. FJ023832), adw1/2 (n=3) and adr (FJ023979). The remaining clones were attributed to ayw1 (n=52) and adr (n=4).

6 Discussion

6.1 Genotypes

The 42 HBsAg positive donors from which clones were derived, were infected with a large variety of subgenotypes (B1, B2, B4, C1, C5, I1, I2) (Table 17) similar to the one observed in our earlier study (220). Both genotypes B and C seemed to circulate with high prevalences in the investigated cohort and were found in 85.7% and 57.1% of investigated donors, with subgenotypes B4 (73.8% of all donors) and C1 (57.1%) dominating. Subgenotyping of a number of genotype B clones was, however, hampered by the mis-assignment of subgenotypes B3, B5, B7 and B8 strains currently available on GenBank (Figure 26). The low phylogenetic support and genetic distances, sometimes below 3% (between B3 and B7) (212), may require a later reclassification of genotype B strains.

The proposed genotype I, reported from several locations in Laos and from Vietnam (106, 220, 248, 294) shows evidence of recombinations between genotypes C, G and

perhaps A (by SimPlot analyses), but is for the purpose of this study not considered a recombination. While genotype I had been proposed on the basis of at least 7.8% mean genetic distance to established genotypes (220), its classification as a genotype remains an open but academic debate (153) until more detailed guidelines for nomenclature and the definition of recombinants are established.

As a result of the co-circulation of HBV subgenotypes in Laos, 5.8% (n=26) of the 446 analysable HBsAg positive blood donors, were mixed infected with multiple HBV variants (Table 16). 17 of these (3.8% of the 446 donors or 65.4% of the 26 mixed infected donors) revealed at least one recombinant strain, all originating from genotypes B and C (subgenotype C1).

The proposed genotype I, however, was relatively rare (3/42 donors). The most prominent Asian recombinants with B and C parent sequences (subgenotypes B2-B4, formerly Ba) were found in Laos in 31/42 donors, while the C/D recombinant predominant in Tibet (67) was absent. These latter widespread recombinant strains (B2-B4, I1, I2) seem to have long evolutionary histories in Asia, originating from either single or multiple very similar recombination events in a distant past, and are here no longer considered recombinants. However, these subgenotypes of B recombined further in Laos with HBV/C, most often C1, strains, to form the relatively recent recombinants described in this study.

6.2 Recombinations

A surprisingly large number of 72 recombinants was found in 64.5% of mixed infected donors, representing 9.9% of all cloned strains. These strains largely reflect recombinant strains circulating in Laos and cannot be explained by PCR artefacts, as confirmed by our control experiments.

127 of the genotypable 185 fragments belonged to the same genotype as the co-infecting, non-recombinant strains. For 60 of these, the most closely related known virus was found, although not necessarily exclusively, in the same donor. Using a published mutation rate of 4.2×10^{-5} mutations per site and year (83), the observed minimum genetic distances of 0–0.7% between parent strain and recombinant fragment would correspond to 0–83 years of evolution from a MRCA. For 13 of these fragments (all originating from different cloned recombinants) even a lower or the same time of evolution than the age of the donor was calculated. Although, some donors with recombinant strains were not co-infected with similar (non-recombinant)

parent strains, the above observations are suggestive of recent recombination events and even within the individual donors. These putative recent recombination events would explain the very limited spread of each of the recombinants and that the recombinant virus strains are largely distinct from each other (Figure 27). Only the proposed genotype I and subgenotypes B2 and B4 represent older recombinants in this region. The large number of apparently independent recent recombination events is also reflected in an unexpected diversity in recombination breakpoints (Figure 27), the locations of which were highly variable throughout the S gene. This variability in recombination sites was surprising, since earlier studies (38, 41, 195, 265) reported recombination breakpoints mostly near gene boundaries. Interestingly, all recombined fragments, independently of their breakpoints within the S gene, were separated by homologous stretches of at least 17 nucleotides, suggesting that this may be conducive to recombinations. The breakpoint diversity in these recombinants also seems to be in contrast to the dominant recombinants found in the rest of mainland Asia (51, 67, 209, 220).

When comparing the breakpoints of individual recombinant clones more closely, clones with different recombination patterns were not only found in different donors, but also within the same donor (e.g. donor M05-1173) and in similar or different strains from different donors (Figure 27). The relatively rare recombinations with identical recombination patterns of the same genotypes in different donors seem to further suggest their recent emergence, probably within the last century. Thus, several of the above lines of evidence, as well as our control experiments, suggest that many of the recombinants described here have occurred in the same donor.

Surprisingly, amino acid substitutions were found in a large proportion of investigated clones. However, these were largely free of vaccine or treatment induced mutations. They furthermore agreed largely with the HBsAg subtypes that had been described previously in genotype B and C strains from Asia (209), indicating that recombination between genotypes B and C is not adding to the diversity of HBsAg subtypes in HBV strains circulating in Laos.

Perinatal transmission is considered by some authors the most common route of infection in Asia (95). Furthermore, superinfections are considered to be rare and the ongoing immune response and the replicative space may limit superinfections and acceleration of chronic HBV even in regions of high endemicity (129). Since furthermore quasispecies found within the individual donors are unlikely to all be

transmitted independently, one could argue that all quasispecies would emerge from the initial pool of viruses infecting at birth. Assuming further, that all quasispecies develop from a single virus, we calculated the mutation rate, required for the most distant quasispecies (of the same subgenotype) to develop since birth of the donors. The median mutation rate was found to be 1.58×10^{-4} mutations per site and year, irrespective of whether samples were included for which at least 5 or 10 clones were available. This is well within the range of published short-term mutation rates (range 1.4×10^{-5} to 7.9×10^{-4}) (327). These estimates of short-term evolutionary rates of HBV are, however, distinct from the much slower mutation rates (~10^{-9}) recently observed during co-evolution of HBV genomes in birds over several millions of years (96). The dramatic differences between long and short-term mutation rates reflect the different evolutionary constraints during co-evolution with the host species and rapid adaption under the pressure of the host immune system. In addition, mutational saturation could be limiting and high mutation rates would not be reflected in the long-term evolution of a virus.

On the basis of the mutation rate 1.58×10^{-4} mutations per site and year, calculated above, the tMRCA would be reduced by a factor of 3.8, largely corresponding to about 0–22 years of evolution for those recombinant fragments for which the most closely related known virus was found within the same donor, further suggesting that these recombinations occurred during the lifetime of the donor. Even under the more likely assumption that more than a single virus quasispecies was infectious at birth, the above contention seems to suggest that the number of quasispecies transmitted vertically is limited.

In conclusion, in Laos, multiple HBV sub-/genotypes co-circulate, generating highly variable recent recombinants with unique breakpoints, few of which seem to have spread within the population. About 65% of mixed infected donors showed recent recombinations in the S-gene alone, strongly suggesting that, at least in Laos, mixed infections invariably lead to recombinations.

Part 7: Occult hepatitis B infections among blood donors in Lao PDR

This manuscript has been accepted for publication as:

Jutavijittum P, Andernach IE, Yousukh A, Samountry B, Samountry K, Thammavong T, Keokhamphue J, Toriyama K and. Muller CP. Occult hepatitis B infections among blood donors in Lao PDR. Vox Sanguinis. In press.

I.E. Andernach contributed to the experimental setup, data analysis and writing of the manuscript.

In Asia, where hepatitis B virus (HBV) is highly endemic, chronically infected individuals with detectable HBV surface antigen (HBsAg) represent a significant risk of virus transmission by blood donation. The Asian genotypes B and C are often found at higher viral loads in the blood (>10^7-10^8 copies/ml) (167), whereas other genotypes, such as genotype E in Africa, are associated with HBeAg seroconversion before the age of 15 or 16 and lower viral loads (48, 167). An additional risk for transfusion transmitted infections (TTI) results from occult HBV infections (OBI), defined as the presence of HBV DNA in the blood or liver without detectable HBsAg. The patho-physiological basis of OBI is still poorly understood but is probably due to an incomplete immune control of the infection. Most individuals with OBI have antibodies against HBV core antigen (anti-HBc) possibly reflecting an ongoing chronic infection with HBsAg below detection limits. This is particularly common in endemic countries where individuals are infected during early childhood. OBI with antibodies against HBV surface antigen (anti-HBs), with or without anti-HBc antibodies, sometimes occurs during the recovery phase of the infection. Rare seronegative cases of OBI occur in individuals with no markers of HBV infection other than HBV DNA. Individuals with chronic hepatitis who have mutated HBsAg that is not detected by diagnostic assays also contribute to OBI. Whilst OBI infections are

generally characterized by very low levels of HBV DNA in the blood (usually <1000 copies/ml or <200 IU/ml (242)), they nevertheless are a concern for recipients of blood donations (2, 169, 214).

In Western countries with a low prevalence of HBV infections, anti-HBc antibody testing is included in the screening of blood donations for TTI to exclude individuals with past exposure to HBV (including the vast majority of OBI). However, in highly endemic countries, this approach would exclude a large proportion of healthy donors that have cleared HBV infections, severely undermining the blood supply (47, 218). In these settings, screening for TTI commonly includes HBsAg testing (169, 218). However, this strategy does not detect OBI or acute infections during the initial infection period. Thus, nucleic acid testing (NAT) has been proposed and tested as a highly sensitive technique for HBV DNA detection in OBI blood donors. However, in low income-countries NAT may not be possible due to limited resources (47, 169, 218, 233).

In Lao PDR, HBsAg is highly endemic, with a prevalence of 8.7% in Lao blood donors (126). As in its neighboring countries Cambodia and Vietnam (218), blood donations in Lao PDR are screened for HBsAg only, with the risk of TTI by HBsAg-negative blood donors with OBI. In this study, we determined the serological profile of Lao blood donors and the prevalence of OBI. We discuss possible implications for the safety of Lao blood donations.

1 Study group

A total of 906 first-time blood donors, including 296 females, were randomly recruited between March and June 2006. In Lao PDR, all blood donors are routinely asked for sexual and medical risk factors related to TTI including clinical hepatitis in an interview and by questionnaire. Donors with previous hepatitis B, C, HIV and syphilis were excluded. Donors are not routinely asked whether they are vaccinated against HBV and sera are not routinely tested for ALT. The mean age of the donors was 23.8 ± 7.4 years (range 16-53 years), with about 80% of donors being less than 30 years old. Serum samples were obtained from Vientiane municipality, Vientiane province and Bolikhamsay province, either at the National Blood Transfusion Centre of the Lao Red Cross in Vientiane or by mobile collecting units and were stored at -20°C until further testing.

2 Serological analyses

All assays were performed on an automated AxSYM® system. HBsAg, HBeAg and anti-HBs, anti-HBc, anti-HBe antibodies were detected by AxSYM® HBsAg V2, AxSYM® HBe 2.0, AxSYM® AUSAB, AxSYM® CORE and AxSYM® Anti-HBe 2.0 assays, respectively.

3 Molecular and phylogenetic analyses

DNA was extracted using the DNA Blood Mini Kit and the HBV genome was amplified as 4 overlapping fragments (preS, S, X, and C). The sensitivity of the PCR reactions for preS, S, X and C were 1424, 1009, 2240 and 950 copies per reaction, respectively. Phylogenetic analyses were performed using MEGA v.5. Genotyping was performed by analyzing at least one of the four fragments. For detection of OBI, DNA was extracted from HBsAg negative, anti-HBc and/or anti-HBs positive donors. HBV DNA was quantified by TaqMan PCR and compared to a HBV plasmid dilution series ranging from 18.5 to $1.85x10^{10}$ copies in the PCR corresponding to 1110 to $111x10^{10}$ copies/ml serum. Although viral DNA can be detected below the linear range, they were not quantifiable. Therefore, values below this linear range (CT values above 34) are expressed as "less than 1110 copies/ml". The estimated limit of detection using our in-house plasmid was about 110 copies/ml or approximately 20 IU/ml. Samples with CT values above 37 were considered negative. Negative controls were included in all extractions and PCR experiments.

4 Statistical analyses

A computerized data sheet was used for record keeping; all data were evaluated with SPSS 17.0 for Windows statistic software package or SigmaPlot 12.0. Parameters were compared by chi-square test or Fisher's exact test where appropriate.

5 Ethical statement

This study was approved by the competent Research Ethics Committee of the Faculty of Medicine of Chiang Mai University, Thailand, and the Ethics Committee of the Faculty of Medical Sciences, National University of Laos, Lao PDR.

6 Results

HBsAg, anti-HBs and anti-HBc antibodies. In the 906 blood donors the overall HBsAg prevalence was 9.6% (87/906), while anti-HBc and anti-HBs antibodies were detected in 43.7% (396/906) and 27.7% (251/906) of donors, respectively. Anti-HBc and anti-HBs antibody prevalence increased with age with significantly higher percentages in the oldest compared to the youngest age group ($p<0.005$ and $p<0.05$, respectively), while HBsAg was less frequent in the older age groups, with a peak prevalence of 11.6% at 20-29 years of age (Figure 28, Table 19). Overall, males were more frequently infected with HBV than females (48.7% and 38.9%, respectively; $p<0.05$). 86 out of the 87 HBsAg-positive donors were also positive for anti-HBc, while one (a 22 year old male) was repeatedly negative for both anti-HBc and anti-HBs, suggestive of recent infection. One of the HBsAg-positive samples was also positive for anti-HBs (Table 19).

494 of the 819 HBsAg-negative sera were negative for all serum markers, while the remaining 325 sera were positive for either anti-HBc antibodies (n=75; 9.2% of HBsAg-negative sera), anti-HBs antibodies (n=15; 1.8%) or both (n=235; 28.7%) (Table 19).

HBeAg and anti-HBe antibodies. 85 HBsAg-positive carriers, excluding two samples with insufficient serum volumes, were further screened for HBeAg and anti-HBe antibodies. HBeAg and anti-HBe antibodies were mutually exclusive and were found in 44.7% (38/85) and 55.3% (47/85) of HBsAg-positive donors, respectively (Table 20). At a younger age (<30 years), approximately half of the carriers were positive for either HBeAg or anti-HBe. In contrast all 10 donors older than 30 years were positive for anti-HBe (Table 20).

6.1 Occult infections

To determine the prevalence of OBI, 320 of the 325 HBsAg-negative, anti-HBc and/or anti-HBs positive sera were screened with a highly sensitive in-house TaqMan PCR. In 35 of the 320 samples (10.9%) HBV DNA was detected, with 32 of these having less than 1110 copies/ml. All 35 TaqMan positive samples were found to be anti-HBc positive, with 23 of these also positive for anti-HBs (Table 21). HBV viral load ranged from below 1110 to 17,550 copies/ml, with the highest concentration observed in anti-HBc positive/anti-HBs negative.

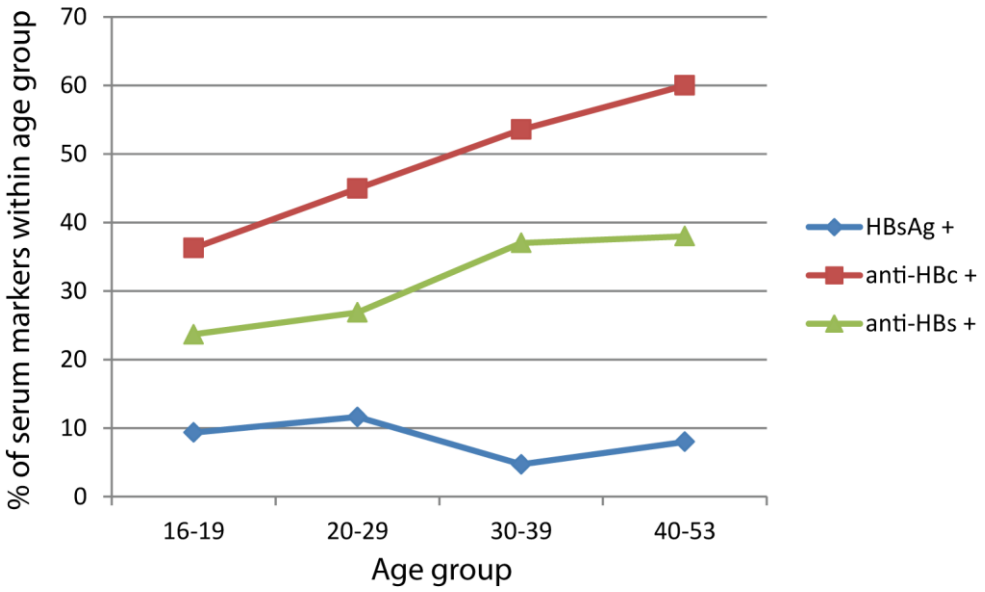

Figure 28: Percentage of HBsAg, anti-HBs and anti-HBc positives among age groups. Prevalence of HBV serum markers according to age group. Prevalences for HBsAg (blue), anti-HBc (red) and anti-HBs (green) are calculated based on the whole study group (906 individuals).

Table 19: Seroprevalence of HBsAg, anti-HBc, and anti-HBs according to age groups.

Serological markers			No. samples (% of age group)				
HBsAg	**Anti-HBc**	**Anti-HBs**	**16-19**	**20-29**	**30-39**	**40-53**	**Total**
+	-	-	0 (0.0%)	1 (0.2%)	0 (0.0%)	0 (0.0%)	1 (0.1%)
+	+	+	1 (0.3%)	0 (0.0%)	0 (0.0%)	0 (0.0%)	1 (0.1%)
+	+	-	31 (9.0%)	44 (11.4%)	6 (4.7%)	4 (8.0%)	85 (9.4%)
-	+	-	17 (5.0%)	29 (7.5%)	20 (15.8%)	9 (18.0%)	75 (8.3%)
-	+	+	75 (21.9%)	101 (26.1%)	42 (33.1%)	17 (34.0%)	235 (25.9%)
-	-	+	5 (1.5%)	3 (0.8%)	5 (3.9%)	2 (4.0%)	15 (1.7%)
-	-	-	213 (62.3%)	209 (54.0%)	54 (42.5%)	18 (36.0%)	494 (54.5%)

Table 20: Seroprevalence of HBeAg and anti-HBe in HBsAg positive sera according to age groups.

Serological markers		No. samples (% of age group)				
HBeAg	Anti-HBe	16-19	20-29	30-39	40-53	Total
+	-	16 (51.6%)	22 (50.0%)	0 (0%)	0 (0%)	38 (44.7%)
-	+	15 (48.4%)	22 (50.0%)	6 (100%)	4 (100%)	47 (55.3%)

Table 21: DNA positivity among HBsAg negative, anti-HBc and/or anti-HBs positive donors.

	Serological markers		No. samples (% of age group)				
DNA	Anti-HBc	Anti-HBs	16-19	20-29	30-39	40-53	Total
+	+	-	3 (3.2%)	3 (2.3%)	5 (7.4%)	1 (3.6%)	12 (3.8%)
+	+	+	10 (10.8%)	8 (6.1%)	2 (2.9%)	3 (10.7%)	23 (7.2%)
+	-	+	0 (0%)	0 (0%)	0 (0%)	0 (0%)	0 (0%)
-	+	-	12 (12.9%)	28 (21.4%)	16 (23.5%)	8 (28.6%)	64 (20.0%)
-	+	+	63 (67.7%)	89 (67.9%)	40 (58.8%)	14 (50%)	206 (64.4%)
-	-	+	5 (5.4%)	3 (2.3%)	5 (7.4%)	2 (7.1%)	15 (4.7%)

Table 22: Prevalence of HBV genotypes B and C according to age groups and HBeAg/anti-HBe status in HBsAg positive sera.

	Serological markers		No. samples (% of age group)				
Genotype	HBeAg	Anti-HBe	16-19	20-29	30-39	40-53	Total
B	+	-	5 (29.4%)	10 (41.7%)	0 (0%)	0 (0%)	15 (34.9%)
	-	+	3 (17.6%)	0 (0%)	1 (50%)	0 (0%)	4 (9.3%)
C	+	-	7 (41.2%)	10 (41.7%)	0 (0%)	0 (0%)	17 (39.5%)
	-	+	2 (11.8%)	1 (4.2%)	1 (50%)	0 (0%)	4 (9.3%)
I	+	-	0 (0%)	1 (4.2%)	0 (0%)	0 (0%)	1 (2.3%)
	-	+	0 (0%)	2 (8.3%)	0 (0%)	0 (0%)	2 (4.7%)

6.2 HBV genotypes

Of the 85 HBsAg positive sera that were screened for HBeAg and anti-HBe, 43 were PCR positive on at least one of the four HBV genotyping fragments. Most HBeAg positive sera (33/38) could be genotyped but as a result of reduced viral load after seroconversion only few anti-HBe positive samples (10/47) were PCR positive and could be genotyped. Phylogenetic analyses revealed that strains belonged to genotypes B (19/43), C (21/43) and I (3/43), with similar rates of genotypes B and C throughout the different age groups (Table 22).

Additionally, the 35 TaqMan PCR positive OBI sera were subjected to preS and S-gene genotyping PCRs. 11 of these, with viral copies ranging from less than 1110 to 17,550 per ml, were found to be positive in at least one of the PCRs. Phylogenetic analyses revealed a predominance of HBV genotype C (8/11), while one strain each was attributed to genotypes B and I. Another strain was suspected to be mixed infected with multiple genotypes and was not further analyzed. While the proportion of genotype C infections in the occult group was significantly higher than in the HBsAg positive group ($p<0.05$), this is likely to be a selection bias, since HBV/C infections tend to have higher viral loads than HBV/B infections. Thus, OBI with HBV/B variants is more likely to remain undetected.

7 Discussion

In healthy first-time blood donors we found a high prevalence of HBV serum markers. We determined that 45.5% of the blood donors were positive for any of the HBV markers, suggesting that almost half of the Lao population has been in contact with HBV at one point during their lives. 9.6% of the donors were HBsAg positive. This is consistent with the 8.7% HBsAg prevalence reported before from Lao PDR (126) as well as from neighboring countries such as Cambodia (7.7% HBsAg and 58.6% anti-HBc prevalence) and Vietnam (11.4% HBsAg and 51.7% anti-HBc) (218).

The presence of anti-HBc and also anti-HBs antibodies in the investigated cohort increased with age and was significantly higher in the oldest age group, while HBsAg-positivity had a tendency to decrease in the higher age groups. This is likely to represent spontaneous HBsAg seroclearance during late stages of HBV infections and is consistent with clearance rates of up to 45% observed in endemic regions (60). In Lao PDR, a country with a high prevalence of chronic HBV carriers, most infections

are acquired during early life, perpetuating chronic infections. Anti-HBs antibodies were found in 27.7% of all donors and only 6.1% of these were anti-HBc negative and therefore potentially vaccinated. This is in accordance with the low HBV vaccine coverage in adults in Lao PDR where HBV vaccine was only introduced in 2001 for infants. Furthermore, 8.3% of donors were anti-HBc-positive but negative for HBsAg and anti-HBs, indicating high numbers of later stage chronic HBV infections with loss or reduced levels of HBsAg, or resolved infections (HBV DNA negative) with anti-HBs antibodies below the detection limit (Figure 28, Table 19).

Interestingly, one sample was repeatedly positive for HBsAg as well as for anti-HBc and anti-HBs antibodies. Whilst this may be a case of exposure to heterologous HBsAg serotypes, we have no further basis for such speculations (7). This serological profile may also be indicative of acute HBV infection or the case of a chronic carrier with severe liver disease, although the latter is unlikely as blood donors are normally healthy volunteers.

More than 40% (38/85) of HBsAg carriers in this study were HBeAg positive. Interestingly, while about half of the carriers below 30 years of age were positive for either HBeAg or anti-HBe, all 10 donors older than 30 years were positive for anti-HBe but not HBeAg (Table 20). This suggests that seroconversion from HBeAg to anti-HBe, an indicator of sustained remission, occurs at around 30 years of age in Lao PDR. Although previous studies indicated earlier HBeAg seroconversion in patients infected with HBV genotype B than for those infected with HBV genotype C (130, 319), we could not observe a difference in the prevalence of HBV B and C genotypes in the different age groups (Table 22).

In our study, 10.9% (35/320) of HBsAg negative, anti-HBV antibody positive donors were HBV DNA positive in a highly sensitive real-time PCR. Thus, at least 3.9% of all 906 investigated donors were occult infected with HBV (OBI). When a less sensitive commercial assay was used, only four samples, including the three with the highest DNA levels, were DNA positive (data not shown). This would correspond to an OBI prevalence of 1.3% of HBsAg negative, anti-HBV antibody positive donors and 0.4% of all donors. It is possible that the individuals detected with the commercial assay represent the greatest risk of onwards HBV transmission. This apparent difference in prevalence highlights the difficulty of comparing OBI prevalence data from studies using different detection techniques. Previous studies in highly endemic regions also showed high prevalences of OBI. In Indonesia, where 9.4% of the population is

chronically infected, a study using nested PCR found 8.1% of blood donors to be occult infected (288). Similarly, in a highly endemic population in India, an in-house PCR detected OBI in 30% of anti-HBc antibody positive donors (228). This is in contrast to data from regions where HBV is not endemic, such as the USA, where OBI can be as low as 3.7% of anti-HBc antibody positive donors (143). While the prevalence of OBI is dependent on epidemiological aspects such as the route of infection, the clinical evolution of the disease and interventions, the infectivity of OBI blood as well as the determinants of this infectivity still remain unclear.

The majority of OBIs that could be genotyped in the current study were attributed to genotype C (8/11). Significantly fewer HBV genotype B infections were observed in the OBI group (1/11) than in HBsAg carriers (19/43), where equal proportions of HBV genotype C and B were observed. One study of OBI in individuals from south-east China, where genotype B is dominant, also showed that the prevalence of genotype C was significantly higher in OBI than in HBsAg positive individuals (318), whilst others have not seen this difference (49). Studies on larger cohorts are needed to further determine the significance of a genotype bias in OBI.

All OBI donors were anti-HBc positive while the prevalence of OBI was somewhat lower in anti-HBs/HBc antibody double positive (23/229; 10%) than in anti-HBc single positive samples (12/76; 15.8%; not statistically significant). In other studies, OBI prevalences were also higher in individuals with anti-HBc antibodies only and lower in double positive or anti-HBs antibody only donors (292). In general, high levels of anti-HBs antibodies in donated blood forestall HBV infections of recipients (reviewed in (47)). Thus, recipient infection may not be prevented in the presence of low levels of anti-HBs antibodies (less than 100 IU/ml), in particular when levels of HBV DNA are high or when the recipients are immuno-compromised (47). In our cohort, OBI sera had a range of anti-HBs levels from 0 to 768 IU/ml. Importantly, the three donors with highest DNA levels (1790, 12800 and 17550 copies per ml) had undetectable anti-HBs antibodies, re-emphasizing the higher risk of TTI from these donors. Rare seronegative cases of OBI occur in individuals with no markers of HBV infection other than HBV DNA. For example, in Thai blood donors about one in 3,000 had HBV DNA and were antibody and HBsAg negative. Approximately 20% of these were window period cases (58, 174). Since we only investigated HBV DNA in antibody positive cases, our study would have missed DNA positive (antibody-negative) window cases.

The significance of OBI for the donor and the recipient is poorly understood. In particular, there are very few studies investigating the infectivity of transfusion with OBI blood. The few available studies seem to agree that only subsets of recipients of OBI blood become infected. For example, one study in Taiwan showed that two out of 11 recipients of OBI blood became subsequently infected (302). In another study from the Japanese Red Cross, OBI blood was infectious in 19% of recipients (251). Recently, it was suggested that the 50% infectious dose of OBI blood products would be about 1000 HBV DNA copies (3). However, the relationship between viral DNA concentration or serology of the donor and infection of the recipient was not determined in these studies. Thus, certainly a large proportion of the OBIs observed in the current study, and especially the "anti-HBc only" donors with higher serum DNA levels, represent a considerable risk of TTI.

To eliminate the risk of HBV transmission by blood transfusion, additional analyses besides HBsAg are required in Lao PDR. While anti-HBc antibody screening would largely exclude the HBsAg negative/DNA positive sera, it would also unnecessarily exclude anti-HBc antibody positive blood with undetectable HBV DNA i.e. approximately 30% of all donors. In Thailand, a country with an emerging economy, blood banks systematically screen for HBV, hepatitis C virus and HIV-1/2 by nucleic acid testing (NAT) since 2006. Together with a broad hepatitis B vaccination strategy, this has led to a reduction in the incidence of new HBV infections in that country (58). Whilst this triple NAT screening may be cost effective in Thailand due to its high prevalence of HIV-1 infections, its cost-benefit in resource poor Lao PDR requires careful evaluation.

Chapter V: Conclusion and Perspectives

Hepatitis B virus (HBV) is a major public health concern. Worldwide, more than 2 billion people are or have been infected with hepatitis B virus at some time during their lives, accounting for approximately 600,000 deaths per year, most of which occur in the developing world. 350 to 400 million people are chronically infected and are at risk of liver cirrhosis and death due to liver failure. Especially in areas of high HBV prevalence, concurrent infections, e.g. with hepatitis D virus (HDV) complicate the course of disease, generally resulting in more severe and progressive liver disease than HBV infection alone.

The first parts of the current dissertation on HBV in sub-Saharan Africa confirm that one of the two genotypes A and E dominates in most countries. We demonstrate that genotype A is more diverse in Africa (4.00% mean genetic diversity) than in the rest of the world (2.96%), suggesting an African origin and a long history on the continent. Genotype E, on the other hand, revealed a conspicuously low genetic diversity (1.75%), indicating a short history of this genotype in Africa. This, however, seems to be in contrast with its excessively high prevalence and its extensive spread throughout the vast West African genotype E crescent. As the spread of pathogens is closely linked to their host, African slaves that were force-migrated to the Americas from the 17th to the early 19th century during the transatlantic slave trade would have disseminated African HBV strains present at that time in the New World. Nevertheless, analysing more than 1600 sequences, we show that genotype E is only sporadically found in the Americas, corroborating an only recent emergence of this genotype.

We furthermore demonstrate that in Haiti, where >90% of the population are descendants of African slaves, more than 40% of HBV infections were caused by genotype A1, which today is found mainly in Eastern Africa. Another 20% belong to a rare subgenotype, A5, which has been found only in the former Bight of Benin, a primary slave trading post. Haitian A subgenotypes appear to have separated early from the African subgenotypes, while the most prevalent genotype and subgenotype in West Africa today (E and A3, respectively) are rare in Haiti. These results provide evidence that the dominant (sub-)genotypes in Africa emerged in the general population only after the end of the transatlantic slave trade, within the last 100 to 200 years, and explain the low genetic diversity of genotype E.

In addition, our analyses show that the prevalence of genotype A and thus HBV in general must have traditionally been much lower in West-Africa than it is today. This is

supported by the current low prevalence of genotype A in the genotype E crescent. In addition, the rare and recent recombination events between genotypes E and A indicate, that the clash between genotypes never occurred in this region. These results, together with today's high prevalence of genotype E in much of Africa, suggest that the high HBV endemicity is a recent phenomenon, probably resulting from well-intended injection mass campaigns of the French and Belgian colonial powers with unsafe needles.

Our phylogenetic analyses also provide evidence, that genotype E emerged in the Northern part of the genotype E crescent before spreading to the South. This is confirmed by phylogeographic analyses, suggesting that genotype E originated from the area of Nigeria, before rapidly spreading throughout sub-Saharan Africa. Furthermore, we estimate a substitution rate of $1.9x10^{-4}$ substitutions per site and year and demonstrate that genotype E emerged only recently, most likely within the last 130 years. Interestingly, viral strains found in Haiti seem to be the result of multiple introductions only in the second half of the 20th century, corroborating the absence of a significant number of genotype E strains in West Africa several centuries ago. Together with the strong increase in the effective number of genotype E infections over time, as revealed by Bayesian skyline plot analyses, our results strongly indicate the recent introduction of HBV/E into the general West-African population after the transatlantic slave-trade had come to an end.

Taken together, the hyperendemicity of genotype E and of HBV in general in Africa is likely the result of dramatic changes in the routes of viral transmission in a relatively recent past. The analyses are however limited by the availability of genotype E full genome sequences and sampling is necessarily fragmented and incomplete. As HBV/E strains have a low genetic variability, an increased number of sequences would likely add to the robustness of analyses in future studies. It is thus important to characterize HBV strains from additional countries to further understand the genetic history of genotype E in sub-Saharan Africa. As furthermore all available HBV/E sequences were sampled within the last 13 years, a deeper calibration using 'older' HBV strains would be desirable. This is however hampered by the lack of older samples, as, unlike for HIV, no older paraffinated tissues are available. Although there is no evidence that genotype E is more infectious than genotype A, a competitive advantage of HBV/E cannot be excluded. Thus, a phenotypic characterisation of genotypes E and A would be desirable and analyses of differences in HBV replication

and virus release as well as in HBsAg production should shed light on differences or similarities in infectivity of these genotypes.

The complexity of HBV infection in sub-Saharan Africa is further increased by concurrent infection with hepatitis D virus (HDV), a satellite virus of HBV. Analysing more than 2000 samples from Burkina Faso, the Central African Republic, Chad and Nigeria, we demonstrate that this virus is an important cause of chronic hepatitis. The prevalences vary widely between countries and cohorts, ranging from 0% to 27.3% in asymptomatic carriers and from 1.3% to 50% in liver patients. Surprisingly, 20.5% of HBsAg-positive children from Burkina Faso were HDV-Ab positive, a prevalence almost 10-times higher than the one observed in the mothers. These children are at high risk to develop severe fulminant or chronic hepatitis.

Vast HDV-Ab differences were furthermore observed in liver patients from Nigeria and the Central African Republic. While in Nigeria a large proportion of liver patients were HBsAg positive, only 1.3% and 6.8% of these were HDV-Ab positive, while 50% of liver patients from the Central African Republic were HDV-Ab positive, indicating that in the Central African Republic, in contrast to Nigeria, HDV super-infection might be the cause of chronic hepatitis with frequent severe liver conditions.

In our study we observed the African clades 5 and 6 solely in Nigeria. However, clade 1 predominated in the analysed cohorts, as more than 70% of strains were assigned to HDV-1, originating from Nigeria, the Central African Republic and the Chad. Our analyses indicate that HDV-1 strains originate from Africa, where largely HBV genotype E prevails. Indeed, more than 50% of genotypable HDV and HBV pairs revealed HDV-1 and HBV/E. Nevertheless, a close co-evolution of hepatitis D virus clade 1 and hepatitis B virus genotype E seems unlikely, since HDV-1 can be found worldwide, while HBV genotype E is largely confined to Africa.

As so far only rare studies have been performed on HDV in sub-Saharan Africa, further analyses are warranted to understand the origin and evolution of hepatitis D virus clades and the association with the different HBV genotypes. However, the analysis of HBV genotypes at concurrent HDV infection is complicated by reduced HBV DNA loads. The characterization of these samples would benefit from high throughput next generation sequencing analyses.

In Lao PDR, about 9% of the population are chronic carriers of HBsAg and multiple subgenotypes of genotypes B, C and I co-circulate. We show that almost 6% of

HBsAg positive rejected blood donors are infected with multiple sub-/genotypes. We furthermore provide evidence, that about 65% of the mixed infected donors showed recombinations in the surface gene alone. These had surprisingly diverse recombination breakpoints and involved the predominant genotypes B and C. Our analyses revealed that recombinant virus strains are largely distinct and vary between and within individual donors.

As perinatal transmission is the most important route of transmission in Asia and superinfection of different HBV variants is considered to be rare, quasispecies found in the individual donors are likely to have developed since birth. Assuming further, that all quasispecies have developed from a single virus, we calculated the substitution rate for the most distant quasispecies and observed 1.58×10^{-4} substitutions per site and year. This is well within published short-term mutation rates and close to the 1.9×10^{-4} substitutions per site and year we observed in our study on genotype E in Africa.

Taken together, our results provide evidence that, at least in Laos, hepatitis B virus mixed infections lead to frequent recombinations, largely within individual donors, that so far have not led to new dominant strains in the population.

As our study was focussed on the HBV surface gene, future studies warrant the analysis of larger regions of the genome. This is, however, complicated by experimental challenges. As full-length genome PCRs seem to have a lower sensitivity than the fragmented surface gene PCRs, such a study would not be performed easily by conventional PCRs. Although next generation sequencing would have the same limitations as other PCR based approaches, the high sequencing throughput, combined with good genome coverage should be considered for further analyses.

Aiming to have a further understanding of the serologic and phylogenetic characteristics of HBV in blood donors from Lao PDR, we show that 45.5% of the blood donors were positive for at least one of the hepatitis B virus markers. Thus, in the absence of vaccination against HBV, our results suggest that almost half of the Laotian population has been in contact with this virus at one point during their lives. We furthermore provide evidence, that blood screening procedures in Lao PDR need to be reconsidered. We show that the screening of donated blood solely for HBsAg leaves a sizable risk of hepatitis B virus transmission by HBsAg-negative occult

infected donors in at least 3.9% of blood donations. These would potentially be infective for hepatitis B. However, screening for other viral markers such as antibodies against the HBV core protein (anti-HBc), as performed in western countries, would exclude about 30% of healthy donors that have cleared HBV infections, thus severely undermining the blood supply in Lao PDR.

Nevertheless, additional analyses besides HBsAg are required in Lao PDR to eliminate the risk of HBV transmission by blood transfusion. Additional studies are necessary to evaluate different screening methods, such as nucleic acid testing. However, the cost-benefit of these screening methods in resource poor Lao PDR requires careful evaluation.

Chapter VI:References

1. **Abdou Chekaraou, M., S. Brichler, W. Mansour, F. Le Gal, A. Garba, P. Deny, and E. Gordien.** 2010. A novel hepatitis B virus (HBV) subgenotype D (D8) strain, resulting from recombination between genotypes D and E, is circulating in Niger along with HBV/E strains. J Gen Virol **91:**1609-20.
2. **Alavian, S. M., S. M. Miri, F. B. Hollinger, and S. M. Jazayeri.** 2012. Occult Hepatitis B (OHB) in Clinical Settings. Hepat Mon **12:**e6126.
3. **Allain, J. P., I. Mihaljevic, M. I. Gonzalez-Fraile, K. Gubbe, L. Holm-Harritshoj, J. M. Garcia, E. Brojer, C. Erikstrup, M. Saniewski, L. Wernish, L. Bianco, H. Ullum, D. Candotti, N. Lelie, W. H. Gerlich, and M. Chudy.** 2013. Infectivity of blood products from donors with occult hepatitis B virus infection. Transfusion.
4. **Alter, H. J., R. H. Purcell, J. L. Gerin, W. T. London, P. M. Kaplan, V. J. McAuliffe, J. Wagner, and P. V. Holland.** 1977. Transmission of hepatitis B to chimpanzees by hepatitis B surface antigen-positive saliva and semen. Infect Immun **16:**928-33.
5. **Alvarado Mora, M. V., C. M. Romano, M. S. Gomes-Gouvea, M. F. Gutierrez, F. J. Carrilho, and J. R. Pinho.** 2010. Molecular epidemiology and genetic diversity of hepatitis B virus genotype E in an isolated Afro-Colombian community. J Gen Virol **91:**501-8.
6. **Andernach, I. E., J. M. Hübschen, and C. P. Muller.** 2009. Hepatitis B virus: the genotype E puzzle. Rev Med Virol **19:**231-40.
7. **Andernach, I. E., P. Jutavijittum, B. Samountry, A. Yousukh, T. Thammavong, J. M. Hubschen, and C. P. Muller.** 2012. A high variability of mixed infections and recent recombinations of hepatitis B virus in Laos. PLoS One **7:**e30245.
8. **Andernach, I. E., C. Nolte, J. W. Pape, and C. P. Muller.** 2009. Slave Trade and Hepatitis B Virus Genotypes and Subgenotypes in Haiti and Africa. Emerg Infect Dis **15:**1222-8.
9. **Arankalle, V. A., S. S. Gandhe, B. J. Borkakoty, A. M. Walimbe, D. Biswas, and J. Mahanta.** 2010. A novel HBV recombinant (genotype I) similar to Vietnam/Laos in a primitive tribe in eastern India. J Viral Hepat **17:**501-10.
10. **Araujo, N. M., F. C. Mello, C. F. Yoshida, C. Niel, and S. A. Gomes.** 2004. High proportion of subgroup A' (genotype A) among Brazilian isolates of Hepatitis B virus. Arch Virol **149:**1383-95.
11. **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.
12. **Arauz-Ruiz, P., H. Norder, K. A. Visona, and L. O. Magnius.** 1997. Genotype F prevails in HBV infected patients of hispanic origin in Central America and may carry the precore stop mutant. J Med Virol **51:**305-12.
13. **Arauz-Ruiz, P., H. Norder, K. A. Visona, and L. O. Magnius.** 1997. Molecular epidemiology of hepatitis B virus in Central America reflected in the genetic variability of the small S gene. J Infect Dis **176:**851-8.
14. **Avellon, A., and J. M. Echevarria.** 2006. Frequency of hepatitis B virus 'a' determinant variants in unselected Spanish chronic carriers. J Med Virol **78:**24-36.
15. **Baha, W., M. M. Ennaji, F. Lazar, M. Melloul, E. El Fahime, A. El Malki, and A. Bennani.** 2012. HBV genotypes prevalence, precore and basal core mutants in Morocco. Infect Genet Evol **12:**1157-62.
16. **Bahri, O., I. Cheikh, N. Hajji, A. Djebbi, N. Maamouri, A. Sadraoui, N. B. Mami, and H. Triki.** 2006. Hepatitis B genotypes, precore and core promoter mutants circulating in Tunisia. J Med Virol **78:**353-7.
17. **Baig, S., A. Siddiqui, R. Chakravarty, and T. Moatter.** 2009. Hepatitis B virus subgenotypes D1 and D3 are prevalent in Pakistan. BMC Res Notes **2:**1.

18. **Baig, S., A. A. Siddiqui, W. Ahmed, H. Qureshi, and A. Arif.** 2007. The association of complex liver disorders with HBV genotypes prevalent in Pakistan. Virol J **4:**128.
19. **Bancroft, W. H., R. Snitbhan, R. M. Scott, M. Tingpalapong, W. T. Watson, P. Tanticharoenyos, J. J. Karwacki, and S. Srimarut.** 1977. Transmission of hepatitis B virus to gibbons by exposure to human saliva containing hepatitis B surface antigen. J Infect Dis **135:**79-85.
20. **Bandelt, H. J., P. Forster, and A. Rohl.** 1999. Median-joining networks for inferring intraspecific phylogenies. Mol Biol Evol **16:**37-48.
21. **Banerjee, A., F. Kurbanov, S. Datta, P. K. Chandra, Y. Tanaka, M. Mizokami, and R. Chakravarty.** 2006. Phylogenetic relatedness and genetic diversity of hepatitis B virus isolates in Eastern India. J Med Virol **78:**1164-74.
22. **Barker, L. F., and R. Murray.** 1972. Relationship of virus dose to incubation time of clinical hepatitis and time of appearance of hepatitis--associated antigen. Am J Med Sci **263:**27-33.
23. **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.
24. **Bartenschlager, R., and H. Schaller.** 1988. The amino-terminal domain of the hepadnaviral P-gene encodes the terminal protein (genome-linked protein) believed to prime reverse transcription. EMBO J **7:**4185-92.
25. **Beck, J., and M. Nassal.** 2007. Hepatitis B virus replication. World J Gastroenterol **13:**48-64.
26. **Beck, J., M. Vogel, and M. Nassal.** 2002. dNTP versus NTP discrimination by phenylalanine 451 in duck hepatitis B virus P protein indicates a common structure of the dNTP-binding pocket with other reverse transcriptases. Nucleic Acids Res **30:**1679-87.
27. **Bekondi, C., C. M. Olinger, N. Boua, A. Talarmin, C. P. Muller, A. Le Faou, and V. Venard.** 2007. Central African Republic is part of the West-African hepatitis B virus genotype E crescent. J Clin Virol **40:**31-7.
28. **Bekondi, C., C. M. Olinger, N. Boua, A. Talarmin, C. P. Muller, A. Le Faou, and V. Venard.** 2007. Central African Republic is part of the West-African hepatitis B virus genotype E crescent. J Clin Virol **40:**31-37.
29. **Bielejec, F., A. Rambaut, M. A. Suchard, and P. Lemey.** 2011. SPREAD: spatial phylogenetic reconstruction of evolutionary dynamics. Bioinformatics **27:**2910-2.
30. **Blakely, T. A., M. N. Bates, M. G. Baker, and M. Tobias.** 1999. Hepatitis B carriage explains the excess rate of hepatocellular carcinoma for Maori, Pacific Island and Asian people compared to Europeans in New Zealand. Int J Epidemiol **28:**204-10.
31. **Blow, J. A., M. J. Turell, A. L. Silverman, and E. D. Walker.** 2001. Stercorarial shedding and transtadial transmission of hepatitis B virus by common bed bugs (Hemiptera: Cimicidae). J Med Entomol **38:**694-700.
32. **Blow, J. A., M. J. Turell, E. D. Walker, and A. L. Silverman.** 2002. Post-bloodmeal diuretic shedding of hepatitis B virus by mosquitoes (Diptera: Culicidae). J Med Entomol **39:**605-12.
33. **Blum, H. E., L. Stowring, A. Figus, C. K. Montgomery, A. T. Haase, and G. N. Vyas.** 1983. Detection of hepatitis B virus DNA in hepatocytes, bile duct epithelium, and vascular elements by in situ hybridization. Proc Natl Acad Sci U S A **80:**6685-8.

34. **Blum, H. E., Z. S. Zhang, E. Galun, F. von Weizsacker, B. Garner, T. J. Liang, and J. R. Wands.** 1992. Hepatitis B virus X protein is not central to the viral life cycle in vitro. J Virol **66:**1223-7.
35. **Blumberg, B. S., B. J. Gerstley, D. A. Hungerford, W. T. London, and A. I. Sutnick.** 1967. A serum antigen (Australia antigen) in Down's syndrome, leukemia, and hepatitis. Ann Intern Med **66:**924-31.
36. **Boesecke, C., and J.-C. Wasmuth.** 2012. Hepatitis B, p. 32-43. *In* S. Mauss, T. Berg, J. Rockstroh, C. Sarrazin, and H. Wedemeyer (ed.), Hepatology - A clinical textbook. Flying Publisher.
37. **Bollyky, P. L., and E. C. Holmes.** 1999. Reconstructing the complex evolutionary history of hepatitis B virus. J Mol Evol **49:**130-41.
38. **Bollyky, P. L., A. Rambaut, P. H. Harvey, and E. C. Holmes.** 1996. Recombination between sequences of hepatitis B virus from different genotypes. J Mol Evol **42:**97-102.
39. **Bosch, F. X., J. Ribes, and J. Borras.** 1999. Epidemiology of primary liver cancer. Semin Liver Dis **19:**271-85.
40. **Bouchard, M. J., L. H. Wang, and R. J. Schneider.** 2001. Calcium signaling by HBx protein in hepatitis B virus DNA replication. Science **294:**2376-8.
41. **Bowyer, S. M., and J. G. Sim.** 2000. Relationships within and between genotypes of hepatitis B virus at points across the genome: footprints of recombination in certain isolates. J Gen Virol **81:**379-92.
42. **Bowyer, S. M., L. van Staden, M. C. Kew, and J. G. Sim.** 1997. A unique segment of the hepatitis B virus group A genotype identified in isolates from South Africa. J Gen Virol **78 (Pt 7):**1719-29.
43. **Brechot, C., F. Jaffredo, D. Lagorce, G. Gerken, K. Meyer zum Buschenfelde, A. Papakonstontinou, S. Hadziyannis, R. Romeo, M. Colombo, J. Rodes, J. Bruix, R. Williams, and N. Naoumov.** 1998. Impact of HBV, HCV and GBV-C/HGV on hepatocellular carcinomas in Europe: results of a European concerted action. J Hepatol **29:**173-83.
44. **Bruss, V.** 2007. Hepatitis B virus morphogenesis. World J Gastroenterol **13:**65-73.
45. **Bruss, V., and W. H. Gerlich.** 1988. Formation of transmembraneous hepatitis B e-antigen by cotranslational in vitro processing of the viral precore protein. Virology **163:**268-75.
46. **Buendia, M. A., P. Paterlini Brechot, P. Tiollais, and C. Brechot.** 2005. Hepatocellular carcinoma: molecular aspects in hepatitis B, p. 269-294. *In* H. C. Thomas, S. M. Lemon, and A. J. Zuckerman (ed.), Viral Hepatitis, 3 ed. Blackwell Publishing Ltd, Oxford, UK.
47. **Candotti, D., and J. P. Allain.** 2009. Transfusion-transmitted hepatitis B virus infection. J Hepatol **51:**798-809.
48. **Candotti, D., K. Danso, and J. P. Allain.** 2007. Maternofetal transmission of hepatitis B virus genotype E in Ghana, west Africa. J Gen Virol **88:**2686-95.
49. **Candotti, D., C. K. Lin, D. Belkhiri, T. Sakuldamrongpanich, S. Biswas, S. Lin, D. Teo, Y. Ayob, and J. P. Allain.** 2012. Occult hepatitis B infection in blood donors from South East Asia: molecular characterisation and potential mechanisms of occurrence. Gut **61:**1744-53.
50. **Candotti, D., O. Opare-Sem, H. Rezvan, F. Sarkodie, and J. P. Allain.** 2006. Molecular and serological characterization of hepatitis B virus in deferred Ghanaian blood donors with and without elevated alanine aminotransferase. J Viral Hepat **13:**715-24.
51. **Cao, Z., X. Bai, X. Guo, Y. Jin, G. Qian, and H. Tu.** 2008. High prevalence of hepatitis B virus pre-S mutation and its association with hepatocellular carcinoma in Qidong, China. Arch Virol **153:**1807-12.

52. **Carman, W. F., M. Jazayeri, A. Basuni, H. C. Thomas, and P. Karayiannis.** 2005. Hepatitis B surface antigen (HBsAg) variants, p. 225-241. *In* H. C. Thomas, S. M. Lemon, and A. J. Zuckerman (ed.), Viral Hepatitis, 3 ed. Blackwell Publishing Ltd, Oxford, UK.
53. **Casey, J. L., T. L. Brown, E. J. Colan, F. S. Wignall, and J. L. Gerin.** 1993. A genotype of hepatitis D virus that occurs in northern South America. Proc Natl Acad Sci U S A **90:**9016-20.
54. **Cavinta, L., J. Sun, A. May, J. Yin, M. von Meltzer, M. Radtke, N. G. Barzaga, G. Cao, and S. Schaefer.** 2009. A new isolate of hepatitis B virus from the Philippines possibly representing a new subgenotype C6. J Med Virol **81:**983-7.
55. **Chan, H. L., S. K. Tsui, C. H. Tse, E. Y. Ng, T. C. Au, L. Yuen, A. Bartholomeusz, K. S. Leung, K. H. Lee, S. Locarnini, and J. J. Sung.** 2005. Epidemiological and virological characteristics of 2 subgroups of hepatitis B virus genotype C. J Infect Dis **191:**2022-32.
56. **Chen, H. S., S. Kaneko, R. Girones, R. W. Anderson, W. E. Hornbuckle, B. C. Tennant, P. J. Cote, J. L. Gerin, R. H. Purcell, and R. H. Miller.** 1993. The woodchuck hepatitis virus X gene is important for establishment of virus infection in woodchucks. J Virol **67:**1218-26.
57. **Chien, R. N.** 2008. Current therapy for hepatitis C or D or immunodeficiency virus concurrent infection with chronic hepatitis B. Hepatol Int **2:**296-303.
58. **Chimparlee, N., S. Oota, S. Phikulsod, P. Tangkijvanich, and Y. Poovorawan.** 2011. Hepatitis B and hepatitis C virus in Thai blood donors. Southeast Asian Journal of Tropical Medicine and Public Health **42**.
59. **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.** 2003. Hepatitis B virus genotypes in the United States: results of a nationwide study. Gastroenterology **125:**444-51.
60. **Chu, C. M., and Y. F. Liaw.** 2007. Chronic hepatitis B virus infection acquired in childhood: special emphasis on prognostic and therapeutic implication of delayed HBeAg seroconversion. J Viral Hepat **14:**147-52.
61. **Chu, C. M., and Y. F. Liaw.** 2007. HBsAg seroclearance in asymptomatic carriers of high endemic areas: appreciably high rates during a long-term follow-up. Hepatology **45:**1187-92.
62. **Chu, C. M., and Y. F. Liaw.** 2013. Natural Course of Chronic Hepatitis B Virus Infection, p. 3-35. *In* C. Shih (ed.), Chronic hepatitis B and C : basic science to clinical applications. World Scientific Publishing Co. Pte. Ltd., Singapore.
63. **Chu, C. M., C. T. Yeh, I. S. Sheen, and Y. F. Liaw.** 2002. Acute hepatitis C virus (HCV) infection in chronic carriers of hepatitis B virus (HBV): the impact of underlying active HBV replication on persistence of HCV infection and antibody responses to HCV. Gut **51:**95-9.
64. **CIA.** 2008. The 2008 World Factbook. November 18, 2008. Accessed at https://www.cia.gov/library/publications/the-world-factbook/index.html. .
65. **CIA.** 2008. The 2008 World Factbook. https://www.cia.gov/library/publications/the-world-factbook/index.html.
66. **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.
67. **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.
68. **Cunha, L., C. Plouzeau, P. Ingrand, J. P. Gudo, I. Ingrand, J. Mondlane, M. Beauchant, and G. Agius.** 2007. Use of replacement blood donors to study the

epidemiology of major blood-borne viruses in the general population of Maputo, Mozambique. J Med Virol **79:**1832-40.

69. **Dandri, M., and J. Petersen.** 2012. HBV Virology, p. 65-84. *In* S. Mauss, T. Berg, J. Rockstroh, C. Sarrazin, and H. Wedemeyer (ed.), Hepatoology - A clinical textbook. Flying Publisher.
70. **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.
71. **De Castro, L., C. Niel, and S. A. Gomes.** 2001. Low frequency of mutations in the core promoter and precore regions of hepatitis B virus in anti-HBe positive Brazilian carriers. BMC Microbiol **1:**10.
72. **Dejean, A., P. Sonigo, S. Wain-Hobson, and P. Tiollais.** 1984. Specific hepatitis B virus integration in hepatocellular carcinoma DNA through a viral 11-base-pair direct repeat. Proc Natl Acad Sci U S A **81:**5350-4.
73. **Deny, P.** 2006. Hepatitis delta virus genetic variability: from genotypes I, II, III to eight major clades? Curr Top Microbiol Immunol **307:**151-71.
74. **Devesa, M., C. L. Loureiro, Y. Rivas, F. Monsalve, N. Cardona, M. C. Duarte, F. Poblete, M. F. Gutierrez, C. Botto, and F. H. Pujol.** 2008. Subgenotype diversity of hepatitis B virus American genotype F in Amerindians from Venezuela and the general population of Colombia. J Med Virol **80:**20-6.
75. **Devesa, M., C. Rodriguez, G. Leon, F. Liprandi, and F. H. Pujol.** 2004. Clade analysis and surface antigen polymorphism of hepatitis B virus American genotypes. J Med Virol **72:**377-84.
76. **Dienstag, J. L.** 2008. Hepatitis B virus infection. N Engl J Med **359:**1486-500.
77. **Drucker, E., P. G. Alcabes, and P. A. Marx.** 2001. The injection century: massive unsterile injections and the emergence of human pathogens. Lancet **358:**1989-92.
78. **Drummond, A. J., and A. Rambaut.** 2007. BEAST: Bayesian evolutionary analysis by sampling trees. BMC Evol Biol **7:**214.
79. **Dumpis, U., E. C. Holmes, M. Mendy, A. Hill, M. Thursz, A. Hall, H. Whittle, and P. Karayiannis.** 2001. Transmission of hepatitis B virus infection in Gambian families revealed by phylogenetic analysis. J Hepatol **35:**99-104.
80. **Dumpis, U., M. Mendy, A. Hill, M. Thursz, A. Hall, H. Whittle, and P. Karayiannis.** 2001. Prevalence of HBV core promoter/precore/core mutations in Gambian chronic carriers. J Med Virol **65:**664-70.
81. **EASL.** 2009. EASL Clinical Practice Guidelines: management of chronic hepatitis B. J Hepatol **50:**227-42.
82. **Elisofon, S. A., and M. M. Jonas.** 2006. Hepatitis B and C in children: current treatment and future strategies. Clin Liver Dis **10:**133-48, vii.
83. **Fares, M. A., and E. C. Holmes.** 2002. A revised evolutionary history of hepatitis B virus (HBV). J Mol Evol **54:**807-14.
84. **Foege, W. H., J. D. Millar, and D. A. Henderson.** 1975. Smallpox eradication in West and Central Africa. Bull World Health Organ **52:**209-22.
85. **Forbi, J. C., G. Vaughan, M. A. Purdy, D. S. Campo, G. L. Xia, L. M. Ganova-Raeva, S. Ramachandran, H. Thai, and Y. E. Khudyakov.** 2010. Epidemic history and evolutionary dynamics of hepatitis B virus infection in two remote communities in rural Nigeria. PLoS One **5:**e11615.
86. **Foupouapouognigni, Y., D. N. Noah, M. T. Sartre, and R. Njouom.** 2011. High prevalence and predominance of hepatitis delta virus genotype 1 infection in cameroon. J Clin Microbiol **49:**1162-4.
87. **Frank, C., M. K. Mohamed, G. T. Strickland, D. Lavanchy, R. R. Arthur, L. S. Magder, T. El Khoby, Y. Abdel-Wahab, E. S. Aly Ohn, W. Anwar, and I. Sallam.**

2000. The role of parenteral antischistosomal therapy in the spread of hepatitis C virus in Egypt. Lancet **355:**887-91.

88. **Fujiwara, K., Y. Tanaka, E. Orito, T. Ohno, T. Kato, K. Sugihara, I. Hasegawa, M. Sakurai, K. Ito, A. Ozasa, Y. Sakamoto, I. Arita, A. El-Gohary, A. Benoit, S. I. Ogoundele-Akplogan, N. Yoshihara, R. Ueda, and M. Mizokami.** 2005. Distribution of HBV genotypes among HBV carriers in Benin:phylogenetic analysis and virological characteristics of HBV genotype E. World J Gastroenterol **11:**6410-5.
89. **Gall, D.** 1954. The chemoprophylaxis of sleeping sickness with the diamidines. Ann Trop Med Parasitol **48:**242-58.
90. **Ganem, D., and A. M. Prince.** 2004. Hepatitis B virus infection--natural history and clinical consequences. N Engl J Med **350:**1118-29.
91. **Ganne-Carrie, N., V. Williams, H. Kaddouri, J. C. Trinchet, S. Dziri-Mendil, C. Alloui, N. A. Hawajri, P. Deny, M. Beaugrand, and E. Gordien.** 2006. Significance of hepatitis B virus genotypes A to E in a cohort of patients with chronic hepatitis B in the Seine Saint Denis District of Paris (France). J Med Virol **78:**335-40.
92. **Geggus, D.** 2001. The French Slave Trade: An Overview. William and Mary Quarterly **LVIII, Number 1:**119-138.
93. **Geggus, D.** 2000. La Traite des Esclaves aux Antilles Francaises à la Fin du 18me Siecle: Quelques Aspects du Marche Local, p. 235-245. *In* S. Marzagalli and H. Bonin (ed.), Négoce, Ports et Océans, XVIe-XXe siècles. Mélanges offerts à paul Butel. Presses Universitaires de Bordeaux, Bordeaux.
94. **Gerin, J. L., B. C. Tennant, A. Ponzetto, R. H. Purcell, and F. J. Tyeryar.** 1983. The woodchuck animal model of hepatitis B-like virus infection and disease. Prog Clin Biol Res **143:**23-8.
95. **Ghendon, Y.** 1987. Perinatal transmission of hepatitis B virus in high-incidence countries. J Virol Methods **17:**69-79.
96. **Gilbert, C., and C. Feschotte.** 2010. Genomic fossils calibrate the long-term evolution of hepadnaviruses. PLoS Biol **8**.
97. **Glebe, D.** 2006. Attachment sites and neutralising epitopes of hepatitis B virus. Minerva Gastroenterol Dietol **52:**3-21.
98. **Glebe, D., and S. Urban.** 2007. Viral and cellular determinants involved in hepadnaviral entry. World J Gastroenterol **13:**22-38.
99. **Goldstein, S. T., M. J. Alter, I. T. Williams, L. A. Moyer, F. N. Judson, K. Mottram, M. Fleenor, P. L. Ryder, and H. S. Margolis.** 2002. Incidence and risk factors for acute hepatitis B in the United States, 1982-1998: implications for vaccination programs. J Infect Dis **185:**713-9.
100. **Gong, Y., E. Yao, and J. E. Tavis.** 2001. Evidence that the RNAseH activity of the duck hepatitis B virus is unable to act on exogenous substrates. BMC Microbiol **1:**12.
101. **Grethe, S., J. O. Heckel, W. Rietschel, and F. T. Hufert.** 2000. Molecular epidemiology of hepatitis B virus variants in nonhuman primates. J Virol **74:**5377-81.
102. **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.
103. **Guirgis, B. S., R. O. Abbas, and H. M. Azzazy.** 2010. Hepatitis B virus genotyping: current methods and clinical implications. Int J Infect Dis **14:**e941-53.
104. **Hadziyannis, S. J.** 2011. Natural history of chronic hepatitis B in Euro-Mediterranean and African countries. J Hepatol **55:**183-91.

105. **Halfon, P., M. Bourliere, S. Pol, Y. Benhamou, D. Ouzan, M. Rotily, H. Khiri, C. Renou, G. Penaranda, D. Saadoun, V. Thibault, J. Serpaggi, M. Varastet, M. H. Tainturier, T. Poynard, and P. Cacoub.** 2006. Multicentre study of hepatitis B virus genotypes in France: correlation with liver fibrosis and hepatitis B e antigen status. J Viral Hepat **13:**329-35.
106. **Hannoun, C., H. Norder, and M. Lindh.** 2000. An aberrant genotype revealed in recombinant hepatitis B virus strains from Vietnam. J Gen Virol **81:**2267-72.
107. **Hannoun, C., A. Soderstrom, G. Norkrans, and M. Lindh.** 2005. Phylogeny of African complete genomes reveals a West African genotype A subtype of hepatitis B virus and relatedness between Somali and Asian A1 sequences. J Gen Virol **86:**2163-7.
108. **Harrison, T. J., G. M. Dusheiko, and A. J. Zuckerman.** 2009. Hepatitis Viruses, p. 199-247. *In* A. J. Zuckerman, J. E. Banatvala, J. R. Pattison, P. D. Griffiths, and B. D. Schoub (ed.), Principles and Practice of Clinical Virology, Fifth Edition ed. John Wiley & Sons Ltd., West Sussex, England.
109. **Hasegawa, I., Y. Tanaka, F. Kurbanov, N. Yoshihara, A. El-Gohary, E. Lyamuya, M. Matee, P. Magessa, K. Fujiwara, A. Ozasa, F. Sugauchi, E. Orito, R. Ueda, and M. Mizokami.** 2006. Molecular epidemiology of hepatitis B virus in the United Republic of Tanzania. J Med Virol **78:**1035-42.
110. **Heller, S., and P. Valencia-Mayoral.** 2007. Treatment of viral hepatitis in children. Arch Med Res **38:**702-10.
111. **Herbert, E. W.** 1975. Smallpox inoculation in Africa. J Afr Hist **16:**539-59.
112. **Huang, C., S. C. Chang, I. C. Yu, Y. G. Tsay, and M. F. Chang.** 2007. Large hepatitis delta antigen is a novel clathrin adaptor-like protein. J Virol **81:**5985-94.
113. **Huang, C. R., and S. J. Lo.** 2010. Evolution and diversity of the human hepatitis d virus genome. Adv Bioinformatics**:**323654.
114. **Huang, Z. M., Q. W. Huang, Y. Q. Qin, C. H. Huang, H. J. Qin, Y. N. Zhou, X. Xu, and C. L. Lu.** 2005. Clinical characteristics and distribution of hepatitis B virus genotypes in Guangxi Zhuang population. World J Gastroenterol **11:**6525-9.
115. **Hübschen, J. M., I. E. Andernach, and C. P. Muller.** 2008. Hepatitis B virus genotype E variability in Africa. J Clin Virol **43:**376-80.
116. **Hübschen, J. M., P. O. Mbah, J. C. Forbi, J. A. Otegbayo, C. M. Olinger, E. Charpentier, and C. P. Muller.** 2011. Detection of a new subgenotype of hepatitis B virus genotype A in Cameroon but not in neighbouring Nigeria. Clin Microbiol Infect **17:**88-94.
117. **Hübschen, J. M., J. Mugabo, C. A. Peltier, J. C. Karasi, A. Sausy, P. Kirpach, V. Arendt, and C. P. Muller.** 2009. Exceptional genetic variability of hepatitis B virus indicates that Rwanda is east of an emerging African genotype E/A1 divide. J Med Virol **81:**435-40.
118. **Hughes, S. A., H. Wedemeyer, and P. M. Harrison.** 2011. Hepatitis delta virus. Lancet **378:**73-85.
119. **Huy, T. T., H. Ushijima, V. X. Quang, K. M. Win, P. Luengrojanakul, K. Kikuchi, T. Sata, and K. Abe.** 2004. Genotype C of hepatitis B virus can be classified into at least two subgroups. J Gen Virol **85:**283-92.
120. **Huy, T. T., H. Ushijima, T. Sata, and K. Abe.** 2006. Genomic characterization of HBV genotype F in Bolivia: genotype F subgenotypes correlate with geographic distribution and T(1858) variant. Arch Virol **151:**589-97.
121. **Imazeki, F., M. Omata, and M. Ohto.** 1991. Complete nucleotide sequence of hepatitis delta virus RNA in Japan. Nucleic Acids Res **19:**5439.

122. **Ivaniushina, V., N. Radjef, M. Alexeeva, E. Gault, S. Semenov, M. Salhi, O. Kiselev, and P. Deny.** 2001. Hepatitis delta virus genotypes I and II cocirculate in an endemic area of Yakutia, Russia. J Gen Virol **82:**2709-18.
123. **Jazayeri, M., A. A. Basuni, N. Sran, R. Gish, G. Cooksley, S. Locarnini, and W. F. Carman.** 2004. HBV core sequence: definition of genotype-specific variability and correlation with geographical origin. J Viral Hepat **11:**488-501.
124. **Jutavijittum, P., Y. Jiviriyawat, A. Yousukh, W. Kunachiwa, and K. Toriyama.** 2006. Genotypes of hepatitis B virus among voluntary blood donors in northern Thailand. Hepatol Res **35:**263-6.
125. **Jutavijittum, P., C. M. Olinger, J. Hübschen, A. Yousukh, B. Samountry, T. Thammavong, K. Toriyama, and C. P. Muller.** 2007. Molecular phylogeny of hepatitis B in Lao PDR reveals multiple subtypes of endemic genotypes, a growing prevalence of recombinations within sub- and genotypes and a potential new genotype, Proceedings of the 2007 International Meeting on the Molecular Biology of Hepatitis B Viruses., Rome.
126. **Jutavijittum, P., A. Yousukh, B. Samountry, K. Samountry, A. Ounavong, T. Thammavong, J. Keokhamphue, and K. Toriyama.** 2007. Seroprevalence of hepatitis B and C virus infections among Lao blood donors. Southeast Asian J Trop Med Public Health **38:**674-9.
127. **Kane, A., J. Lloyd, M. Zaffran, L. Simonsen, and M. Kane.** 1999. Transmission of hepatitis B, hepatitis C and human immunodeficiency viruses through unsafe injections in the developing world: model-based regional estimates. Bull World Health Organ **77:**801-7.
128. **Kann, M., and W. H. Gerlich.** 2005. Structure and molecular virology, p. 149-180. *In* H. C. Thomas, S. M. Lemon, and A. J. Zuckerman (ed.), Viral Hepatitis. Blackwell Publishing Ltd, Oxford, UK.
129. **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.
130. **Kao, J. H., P. J. Chen, M. Y. Lai, and D. S. Chen.** 2004. Hepatitis B virus genotypes and spontaneous hepatitis B e antigen seroconversion in Taiwanese hepatitis B carriers. J Med Virol **72:**363-9.
131. **Kaplan, P. M., R. L. Greenman, J. L. Gerin, R. H. Purcell, and W. S. Robinson.** 1973. DNA polymerase associated with human hepatitis B antigen. J Virol **12:**995-1005.
132. **Karasch, M.** 1998. Brazil: An Overview, p. 116-125. *In* P. Finkelman and J. C. Miller (ed.), Macmillan Encyclopedia of World Slavery, vol. 1. Simon & Schuster Macmillan, New York, NY.
133. **Kato, H., K. Fujiwara, R. G. Gish, H. Sakugawa, H. Yoshizawa, F. Sugauchi, E. Orito, R. Ueda, Y. Tanaka, T. Kato, Y. Miyakawa, and M. Mizokami.** 2005. Classifying genotype F of hepatitis B virus into F1 and F2 subtypes. World J Gastroenterol **11:**6295-304.
134. **Kato, H., R. G. Gish, N. Bzowej, M. Newsom, F. Sugauchi, Y. Tanaka, T. Kato, E. Orito, S. Usuda, R. Ueda, Y. Miyakawa, and M. Mizokami.** 2004. Eight genotypes (A-H) of hepatitis B virus infecting patients from San Francisco and their demographic, clinical, and virological characteristics. J Med Virol **73:**516-21.
135. **Katoh, K., G. Asimenos, and H. Toh.** 2009. Multiple alignment of DNA sequences with MAFFT. Methods Mol Biol **537:**39-64.
136. **Katoh, K., and H. Toh.** 2008. Recent developments in the MAFFT multiple sequence alignment program. Brief Bioinform **9:**286-98.

137. **Keasler, V. V., A. J. Hodgson, C. R. Madden, and B. L. Slagle.** 2007. Enhancement of hepatitis B virus replication by the regulatory X protein in vitro and in vivo. J Virol **81:**2656-62.
138. **Keasler, V. V., A. J. Hodgson, C. R. Madden, and B. L. Slagle.** 2009. Hepatitis B virus HBx protein localized to the nucleus restores HBx-deficient virus replication in HepG2 cells and in vivo in hydrodynamically-injected mice. Virology **390:**122-9.
139. **Kedda, M. A., A. Kramvis, M. C. Kew, G. Lecatsas, A. C. Paterson, S. Aspinall, J. H. Stark, W. A. De Klerk, and B. Gridelli.** 2000. Susceptibility of chacma baboons (Papio ursinus orientalis) to infection by hepatitis B virus. Transplantation **69:**1429-34.
140. **Kelly, D., and E. Boxall.** 2005. Neonatal and paediatric infection, p. 714-739. *In* H. C. Thomas, S. M. Lemon, and A. J. Zuckerman (ed.), Viral Hepatitis, 3 ed. Blackwell Publishing Ltd, Oxford, UK.
141. **Kidd-Ljunggren, K., Y. Miyakawa, and A. H. Kidd.** 2002. Genetic variability in hepatitis B viruses. J Gen Virol **83:**1267-80.
142. **Kimbi, G. C., A. Kramvis, and M. C. Kew.** 2004. Distinctive sequence characteristics of subgenotype A1 isolates of hepatitis B virus from South Africa. J Gen Virol **85:**1211-20.
143. **Kleinman, S. H., M. C. Kuhns, D. S. Todd, S. A. Glynn, A. McNamara, A. DiMarco, M. P. Busch, and S. Retrovirus Epidemiology Donor.** 2003. Frequency of HBV DNA detection in US blood donors testing positive for the presence of anti-HBc: implications for transfusion transmission and donor screening. Transfusion **43:**696-704.
144. **Komas, N. P., S. Bai-Sepou, A. Manirakiza, J. Leal, A. Bere, and A. Le Faou.** 2010. The prevalence of hepatitis B virus markers in a cohort of students in Bangui, Central African Republic. BMC Infect Dis **10:**226.
145. **Kramvis, A., K. Arakawa, M. C. Yu, R. Nogueira, D. O. Stram, and M. C. Kew.** 2008. Relationship of serological subtype, basic core promoter and precore mutations to genotypes/subgenotypes of hepatitis B virus. J Med Virol **80:**27-46.
146. **Kramvis, A., M. Kew, and G. Francois.** 2005. Hepatitis B virus genotypes. Vaccine **23:**2409-23.
147. **Kramvis, A., and M. C. Kew.** 2007. Epidemiology of hepatitis B virus in Africa, its genotypes and clinical associations of genotypes. Hepatol Res **37:**S9-S19.
148. **Kramvis, A., and M. C. Kew.** 2007. Molecular characterization of subgenotype A1 (subgroup Aa) of hepatitis B virus. Hepatol Res **37:**S27-32.
149. **Kramvis, A., K. Restorp, H. Norder, J. F. Botha, L. O. Magnius, and M. C. Kew.** 2005. Full genome analysis of hepatitis B virus genotype E strains from South-Western Africa and madagascar reveals low genetic variability. J Med Virol **77:**47-52.
150. **Kramvis, A., L. Weitzmann, W. K. Owiredu, and M. C. Kew.** 2002. Analysis of the complete genome of subgroup A' hepatitis B virus isolates from South Africa. J Gen Virol **83:**835-9.
151. **Kumar, S., K. Tamura, and M. Nei.** 2004. MEGA3: Integrated software for Molecular Evolutionary Genetics Analysis and sequence alignment. Brief Bioinform **5:**150-63.
152. **Kurbanov, F., Y. Tanaka, K. Fujiwara, F. Sugauchi, D. Mbanya, L. Zekeng, N. Ndembi, C. Ngansop, L. Kaptue, T. Miura, E. Ido, M. Hayami, H. Ichimura, and M. Mizokami.** 2005. A new subtype (subgenotype) Ac (A3) of hepatitis B virus and recombination between genotypes A and E in Cameroon. J Gen Virol **86:**2047-56.

153. **Kurbanov, F., Y. Tanaka, A. Kramvis, P. Simmonds, and M. Mizokami.** 2008. When should "I" consider a new hepatitis B virus genotype? J Virol **82:**8241-2.
154. **Lambert, C., T. Doring, and R. Prange.** 2007. Hepatitis B virus maturation is sensitive to functional inhibition of ESCRT-III, Vps4, and gamma 2-adaptin. J Virol **81:**9050-60.
155. **Landers, T. A., H. B. Greenberg, and W. S. Robinson.** 1977. Structure of hepatitis B Dane particle DNA and nature of the endogenous DNA polymerase reaction. J Virol **23:**368-76.
156. **Lavanchy, D.** 2005. Epidemiology, p. 181-192. *In* H. C. Thomas, S. M. Lemon, and A. J. Zuckerman (ed.), Viral Hepatitis, 3 ed. Blackwell Publishing Ltd, Oxford, UK.
157. **Le Gal, F., S. Badur, N. A. Hawajri, F. Akyuz, S. Kaymakoglu, S. Brichler, F. Zoulim, E. Gordien, E. Gault, and P. Deny.** 2012. Current hepatitis delta virus type 1 (HDV1) infections in central and eastern Turkey indicate a wide genetic diversity that is probably linked to different HDV1 origins. Arch Virol **157:**647-59.
158. **Le Gal, F., E. Gault, M. P. Ripault, J. Serpaggi, J. C. Trinchet, E. Gordien, and P. Deny.** 2006. Eighth major clade for hepatitis delta virus. Emerg Infect Dis **12:**1447-50.
159. **Le Gal, F., E. Gordien, D. Affolabi, T. Hanslik, C. Alloui, P. Deny, and E. Gault.** 2005. Quantification of hepatitis delta virus RNA in serum by consensus real-time PCR indicates different patterns of virological response to interferon therapy in chronically infected patients. J Clin Microbiol **43:**2363-9.
160. **Lee, C. H., S. C. Chang, C. H. Wu, and M. F. Chang.** 2001. A novel chromosome region maintenance 1-independent nuclear export signal of the large form of hepatitis delta antigen that is required for the viral assembly. J Biol Chem **276:**8142-8.
161. **Leupin, O., S. Bontron, C. Schaeffer, and M. Strubin.** 2005. Hepatitis B virus X protein stimulates viral genome replication via a DDB1-dependent pathway distinct from that leading to cell death. J Virol **79:**4238-45.
162. **Li, D. D., L. Ding, J. Wang, Q. C. Meilang, X. J. Lu, X. B. Song, C. M. Tao, B. W. Ying, and L. L. Wang.** 2010. Prevalence of hepatitis B virus genotypes and their relationship to clinical laboratory outcomes in Tibetan and Han Chinese. J Int Med Res **38:**195-201.
163. **Liaw, Y. F.** 1995. Role of hepatitis C virus in dual and triple hepatitis virus infection. Hepatology **22:**1101-8.
164. **Liaw, Y. F., Y. C. Chen, I. S. Sheen, R. N. Chien, C. T. Yeh, and C. M. Chu.** 2004. Impact of acute hepatitis C virus superinfection in patients with chronic hepatitis B virus infection. Gastroenterology **126:**1024-9.
165. **Liaw, Y. F., and C. M. Chu.** 2009. Hepatitis B virus infection. Lancet **373:**582-92.
166. **Liaw, Y. F., N. Leung, J. H. Kao, T. Piratvisuth, E. Gane, K. H. Han, R. Guan, G. K. Lau, and S. Locarnini.** 2008. Asian-Pacific consensus statement on the management of chronic hepatitis B: a 2008 update. Hepatol Int **2:**263-83.
167. **Liaw, Y. F., N. Leung, J. H. Kao, T. Piratvisuth, E. Gane, K. H. Han, R. Guan, G. K. Lau, S. Locarnini, and B. G. W. P. o. t. A.-P. A. f. t. S. o. t. L. Chronic Hepatitis.** 2008. Asian-Pacific consensus statement on the management of chronic hepatitis B: a 2008 update. Hepatol Int **2:**263-83.
168. **Lin, C. L., C. J. Liu, P. J. Chen, M. Y. Lai, D. S. Chen, and J. H. Kao.** 2007. High prevalence of occult hepatitis B virus infection in Taiwanese intravenous drug users. J Med Virol **79:**1674-8.
169. **Liu, C. J., D. S. Chen, and P. J. Chen.** 2006. Epidemiology of HBV infection in Asian blood donors: emphasis on occult HBV infection and the role of NAT. J Clin Virol **36 Suppl 1:**S33-44.

170. **Liu, H. F., E. Sokal, and P. Goubau.** 2001. Wide variety of genotypes and geographic origins of hepatitis B virus in Belgian children. J Pediatr Gastroenterol Nutr **32:**274-7.
171. **Liu, Y., M. Hussain, S. Wong, S. K. Fung, H. J. Yim, and A. S. Lok.** 2007. A genotype-independent real-time PCR assay for quantification of hepatitis B virus DNA. J Clin Microbiol **45:**553-8.
172. **Locarnini, S. A., and L. Yuen.** 2010. Molecular genesis of drug-resistant and vaccine-escape HBV mutants. Antivir Ther **15:**451-61.
173. **Lole, K. S., R. C. Bollinger, R. S. Paranjape, D. Gadkari, S. S. Kulkarni, N. G. Novak, R. Ingersoll, H. W. Sheppard, and S. C. Ray.** 1999. Full-length human immunodeficiency virus type 1 genomes from subtype C-infected seroconverters in India, with evidence of intersubtype recombination. J Virol **73:**152-60.
174. **Louisirirotchanakul, S., S. Oota, K. Khuponsarb, W. Chalermchan, S. Phikulsod, V. Chongkolwatana, T. Sakuldamrongpanish, P. Kitpoka, P. Chielsilp, S. Tanprasert, T. Tirawatnapong, C. Wasi, and N. A. T. S. i. T. B. D. Working Group for.** 2011. Occult hepatitis B virus infection in Thai blood donors. Transfusion **51:**1532-40.
175. **Lucifora, J., S. Arzberger, D. Durantel, L. Belloni, M. Strubin, M. Levrero, F. Zoulim, O. Hantz, and U. Protzer.** 2011. Hepatitis B virus X protein is essential to initiate and maintain virus replication after infection. J Hepatol **55:**996-1003.
176. **Lusida, M. I., V. E. Nugrahaputra, Soetjipto, R. Handajani, M. Nagano-Fujii, M. Sasayama, T. Utsumi, and H. Hotta.** 2008. Novel Subgenotypes of Hepatitis B Virus Genotypes C and D in Papua, Indonesia. J Clin Microbiol.
177. **Maïga, I., A. Le Faou, C. P. Muller, and V. Venard.** 2005. Unexpected high prevalence of hepatitis B and HIV infections in Malian medical students. Eur J Clin Microbiol Infect Dis **24:**501-2.
178. **Makuwa, M., A. Mintsa-Ndong, S. Souquiere, D. Nkoghe, E. M. Leroy, and M. Kazanji.** 2009. Prevalence and molecular diversity of hepatitis B virus and hepatitis delta virus in urban and rural populations in northern Gabon in central Africa. J Clin Microbiol **47:**2265-8.
179. **Makuwa, M., S. Souquiere, O. Bourry, P. Rouquet, P. Telfer, P. Mauclere, M. Kazanji, P. Roques, and F. Simon.** 2007. Complete-genome analysis of hepatitis B virus from wild-born chimpanzees in central Africa demonstrates a strain-specific geographical cluster. J Gen Virol **88:**2679-85.
180. **Makuwa, M., S. Souquiere, P. Telfer, C. Apetrei, M. Vray, I. Bedjabaga, A. Mouinga-Ondeme, R. Onanga, P. A. Marx, M. Kazanji, P. Roques, and F. Simon.** 2006. Identification of hepatitis B virus subgenotype A3 in rural Gabon. J Med Virol **78:**1175-84.
181. **Mansour, W., M. A. Bollahi, C. T. Hamed, S. Brichler, F. Le Gal, A. Ducancelle, B. Lo, E. Gordien, M. Rosenheim, and F. Lunel.** 2012. Virological and epidemiological features of hepatitis delta infection among blood donors in Nouakchott, Mauritania. J Clin Virol.
182. **Martin, D., and E. Rybicki.** 2000. RDP: detection of recombination amongst aligned sequences. Bioinformatics **16:**562-3.
183. **Mason, W. S., M. S. Halpern, J. M. England, G. Seal, J. Egan, L. Coates, C. Aldrich, and J. Summers.** 1983. Experimental transmission of duck hepatitis B virus. Virology **131:**375-84.
184. **Mathet, V. L., M. L. Cuestas, V. Ruiz, M. L. Minassian, C. Rivero, J. Trinks, G. Daleoso, L. M. Leon, A. Sala, B. Libellara, D. Corach, and J. R. Oubina.** 2006. Detection of hepatitis B virus (HBV) genotype E carried--even in the presence of high titers of anti-HBs antibodies--by an Argentinean patient of African

descent who had received vaccination against HBV. J Clin Microbiol **44:**3435-9.

185. **Mathet, V. L., M. Feld, L. Espinola, D. O. Sanchez, V. Ruiz, O. Mando, G. Carballal, J. F. Quarleri, F. D'Mello, C. R. Howard, and J. R. Oubina.** 2003. Hepatitis B virus S gene mutants in a patient with chronic active hepatitis with circulating Anti-HBs antibodies. J Med Virol **69:**18-26.
186. **Mbayed, V. A., J. L. Lopez, P. F. Telenta, G. Palacios, I. Badia, A. Ferro, C. Galoppo, and R. Campos.** 1998. Distribution of hepatitis B virus genotypes in two different pediatric populations from Argentina. J Clin Microbiol **36:**3362-5.
187. **McMahon, B. J.** 2009. The influence of hepatitis B virus genotype and subgenotype on the natural history of chronic hepatitis B. Hepatol Int **3:**334-42.
188. **Meier, A., S. Mehrle, T. S. Weiss, W. Mier, and S. Urban.** 2013. Myristoylated PreS1-domain of the hepatitis B virus L-protein mediates specific binding to differentiated hepatocytes. Hepatology.
189. **Meldal, B. H., N. M. Moula, I. H. Barnes, K. Boukef, and J. P. Allain.** 2009. A novel hepatitis B virus subgenotype, D7, in Tunisian blood donors. J Gen Virol **90:**1622-8.
190. **Melegari, M., S. K. Wolf, and R. J. Schneider.** 2005. Hepatitis B virus DNA replication is coordinated by core protein serine phosphorylation and HBx expression. J Virol **79:**9810-20.
191. **Menne, S., and P. J. Cote.** 2007. The woodchuck as an animal model for pathogenesis and therapy of chronic hepatitis B virus infection. World J Gastroenterol **13:**104-24.
192. **Milne, I., D. Lindner, M. Bayer, D. Husmeier, G. McGuire, D. F. Marshall, and F. Wright.** 2009. TOPALi v2: a rich graphical interface for evolutionary analyses of multiple alignments on HPC clusters and multi-core desktops. Bioinformatics **25:**126-7.
193. **Moraes, M. T., S. A. Gomes, and C. Niel.** 1996. Sequence analysis of pre-S/S gene of hepatitis B virus strains of genotypes A, D, and F isolated in Brazil. Arch Virol **141:**1767-73.
194. **Morgan, P. D.** 1998. Slave Trade: Transatlantic, p. 837-844. *In* P. Finkelman and J. C. Miller (ed.), Macmillan Encyclopedia of World Slavery, vol. 2. Simon & Schuster Macmillan, New York, NY.
195. **Morozov, V., M. Pisareva, and M. Groudinin.** 2000. Homologous recombination between different genotypes of hepatitis B virus. Gene **260:**55-65.
196. **Motta-Castro, A. R., R. M. Martins, N. M. Araujo, C. Niel, G. B. Facholi, B. V. Lago, F. C. Mello, and S. A. Gomes.** 2008. Molecular epidemiology of hepatitis B virus in an isolated Afro-Brazilian community. Arch Virol **153:**2197-205.
197. **Mulders, M. N., V. Venard, M. Njayou, A. P. Edorh, A. O. Bola Oyefolu, M. O. Kehinde, J. J. Muyembe Tamfum, Y. K. Nebie, I. Maiga, W. Ammerlaan, F. Fack, S. A. Omilabu, A. Le Faou, and C. P. Muller.** 2004. Low Genetic Diversity despite Hyperendemicity of Hepatitis B Virus Genotype E throughout West Africa. J Infect Dis **190:**400-8.
198. **Mulyanto, S. N. Depamede, K. Surayah, A. A. Tjahyono, Jirintai, S. Nagashima, M. Takahashi, and H. Okamoto.** 2010. Identification and characterization of novel hepatitis B virus subgenotype C10 in Nusa Tenggara, Indonesia. Arch Virol.
199. **Mulyanto, S. N. Depamede, K. Surayah, F. Tsuda, K. Ichiyama, M. Takahashi, and H. Okamoto.** 2009. A nationwide molecular epidemiological study on hepatitis B virus in Indonesia: identification of two novel subgenotypes, B8 and C7. Arch Virol **154:**1047-59.

200. **Mulyanto, S. N. Depamede, A. Wahyono, Jirintai, S. Nagashima, M. Takahashi, and H. Okamoto.** 2011. Analysis of the full-length genomes of novel hepatitis B virus subgenotypes C11 and C12 in Papua, Indonesia. J Med Virol **83:**54-64.
201. **Mulyanto, P. Pancawardani, S. N. Depamede, A. Wahyono, S. Jirintai, S. Nagashima, M. Takahashi, T. Nishizawa, and H. Okamoto.** 2012. Identification of four novel subgenotypes (C13-C16) and two inter-genotypic recombinants (C12/G and C13/B3) of hepatitis B virus in Papua province, Indonesia. Virus Res **163:**129-40.
202. **Nagasaki, F., H. Niitsuma, J. G. Cervantes, M. Chiba, S. Hong, T. Ojima, Y. Ueno, E. Bondoc, K. Kobayashi, M. Ishii, and T. Shimosegawa.** 2006. Analysis of the entire nucleotide sequence of hepatitis B virus genotype B in the Philippines reveals a new subgenotype of genotype B. J Gen Virol **87:**1175-80.
203. **Nakano, T., L. Lu, X. Hu, M. Mizokami, E. Orito, C. Shapiro, S. Hadler, and B. Robertson.** 2001. Characterization of hepatitis B virus genotypes among Yucpa Indians in Venezuela. J Gen Virol **82:**359-65.
204. **Nassal, M.** 2008. Hepatitis B viruses: reverse transcription a different way. Virus Res **134:**235-49.
205. **Naumann, H., S. Schaefer, C. F. Yoshida, A. M. Gaspar, R. Repp, and W. H. Gerlich.** 1993. Identification of a new hepatitis B virus (HBV) genotype from Brazil that expresses HBV surface antigen subtype adw4. J Gen Virol **74 (Pt 8):**1627-32.
206. **Nguyen, V. T., M. L. McLaws, and G. J. Dore.** 2007. Highly endemic hepatitis B infection in rural Vietnam. J Gastroenterol Hepatol **22:**2093-100.
207. **Ni, Y., J. Sonnabend, S. Seitz, and S. Urban.** 2010. The pre-s2 domain of the hepatitis B virus is dispensable for infectivity but serves a spacer function for L-protein-connected virus assembly. J Virol **84:**3879-88.
208. **Nishioka, S., and T. W. Gyorkos.** 2001. Tattoos as risk factors for transfusion-transmitted diseases. Int J Infect Dis **5:**27-34.
209. **Norder, H., A. M. Courouce, P. Coursaget, J. M. Echevarria, S. D. Lee, I. K. Mushahwar, B. H. Robertson, S. Locarnini, and L. O. Magnius.** 2004. Genetic diversity of hepatitis B virus strains derived worldwide: genotypes, subgenotypes, and HBsAg subtypes. Intervirology **47:**289-309.
210. **Norder, H., B. Hammas, S. D. Lee, K. Bile, A. M. Courouce, I. K. Mushahwar, and L. O. Magnius.** 1993. Genetic relatedness of hepatitis B viral strains of diverse geographical origin and natural variations in the primary structure of the surface antigen. J Gen Virol **74 (Pt 7):**1341-8.
211. **Norder, H., B. Hammas, S. Lofdahl, A. M. Courouce, and L. O. Magnius.** 1992. 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 (Pt 5):**1201-8.
212. **Nurainy, N., D. H. Muljono, H. Sudoyo, and S. Marzuki.** 2008. Genetic study of hepatitis B virus in Indonesia reveals a new subgenotype of genotype B in east Nusa Tenggara. Arch Virol.
213. **Nwokediuko, S. C., and U. Ijeoma.** 2009. Seroprevalence of antibody to HDV in Nigerians with hepatitis B virus-related liver diseases. Niger J Clin Pract **12:**439-42.
214. **Ocana, S., M. L. Casas, I. Buhigas, and J. L. Lledo.** 2011. Diagnostic strategy for occult hepatitis B virus infection. World J Gastroenterol **17:**1553-7.
215. **Odemuyiwa, S. O., M. N. Mulders, O. I. Oyedele, S. O. Ola, G. N. Odaibo, D. O. Olaleye, and C. P. Muller.** 2001. Phylogenetic analysis of new hepatitis B virus isolates from Nigeria supports endemicity of genotype E in West Africa. J Med Virol **65:**463-9.

216. **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.
217. **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 (Pt 10):**2575-83.
218. **Ol, H. S., B. Bjoerkvoll, S. Sothy, Y. Van Heng, H. Hoel, A. Husebekk, T. Gutteberg, S. Larsen, and H. Husum.** 2009. Prevalence of hepatitis B and hepatitis C virus infections in potential blood donors in rural Cambodia. Southeast Asian J Trop Med Public Health **40:**963-71.
219. **Olinger, C. M., P. Jutavijittum, J. M. Hubschen, A. Yousukh, B. Samountry, T. Thammavong, K. Toriyama, and C. P. Muller.** 2008. Possible new hepatitis B virus genotype, southeast Asia. Emerg Infect Dis **14:**1777-80.
220. **Olinger, C. M., P. Jutavijittum, J. M. Hübschen, A. Yousukh, B. Samountry, T. Thammavong, K. Toriyama, and C. P. Muller.** 2008. Possible new hepatitis B virus genotype, southeast Asia. Emerg Infect Dis **14:**1777-80.
221. **Olinger, C. M., V. Venard, M. Njayou, A. O. Oyefolu, I. Maiga, A. J. Kemp, S. A. Omilabu, A. le Faou, and C. P. Muller.** 2006. Phylogenetic analysis of the precore/core gene of hepatitis B virus genotypes E and A in West Africa: new subtypes, mixed infections and recombinations. J Gen Virol **87:**1163-73.
222. **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.
223. **Osiowy, C., E. Giles, Y. Tanaka, M. Mizokami, and G. Y. Minuk.** 2006. Molecular evolution of hepatitis B virus over 25 years. J Virol **80:**10307-14.
224. **Owiredu, W. K., A. Kramvis, and M. C. Kew.** 2001. Molecular analysis of hepatitis B virus genomes isolated from black African patients with fulminant hepatitis B. J Med Virol **65:**485-92.
225. **Padidam, M., S. Sawyer, and C. M. Fauquet.** 1999. Possible emergence of new geminiviruses by frequent recombination. Virology **265:**218-25.
226. **Palumbo, E., G. Scotto, G. Faleo, D. C. Cibelli, A. Saracino, and G. Angarano.** 2007. Prevalence of HBV-genotypes in immigrants affected by HBV-related chronic active hepatitis. Arq Gastroenterol **44:**54-7.
227. **Panigrahi, R., A. Biswas, A. Banerjee, S. P. Singh, M. K. Panigrahi, A. M. Roque-Afonso, H. S. Das, P. K. Mahapatra, S. Chakrabarti, and R. Chakravarty.** 2012. Subgenotype D5, BCP and MHR mutations in hepatic complications among hepatitis B virus infected patients from Orissa, India. Infect Genet Evol **12:**1622-9.
228. **Panigrahi, R., A. Biswas, S. Datta, A. Banerjee, P. K. Chandra, P. K. Mahapatra, B. Patnaik, S. Chakrabarti, and R. Chakravarty.** 2010. Anti-hepatitis B core antigen testing with detection and characterization of occult hepatitis B virus by an in-house nucleic acid testing among blood donors in Behrampur, Ganjam, Orissa in southeastern India: implications for transfusion. Virol J **7:**204.
229. **Paraskevis, D., G. Magiorkinis, E. Magiorkinis, S. Y. Ho, R. Belshaw, J. P. Allain, and A. Hatzakis.** 2013. Dating the origin and dispersal of hepatitis B virus infection in humans and primates. Hepatology **57:**908-16.
230. **Pawlotsky, J. M., L. Belec, G. Gresenguet, L. Deforges, M. Bouvier, J. Duval, and D. Dhumeaux.** 1995. High prevalence of hepatitis B, C, and E markers in young sexually active adults from the Central African Republic. J Med Virol **46:**269-72.

231. **Pawlotsky, J. M., L. Tsakiris, F. Roudot-Thoraval, C. Pellet, L. Stuyver, J. Duval, and D. Dhumeaux.** 1995. Relationship between hepatitis C virus genotypes and sources of infection in patients with chronic hepatitis C. J Infect Dis **171:**1607-10.
232. **Pepin, J., and A. C. Labbe.** 2008. Noble goals, unforeseen consequences: control of tropical diseases in colonial Central Africa and the iatrogenic transmission of blood-borne viruses. Trop Med Int Health **13:**744-53.
233. **Phikulsod, S., S. Oota, T. Tirawatnapong, T. Sakuldamrongpanich, W. Chalermchan, S. Louisirirotchanakul, S. Tanprasert, V. Chongkolwatana, P. Kitpoka, P. Phanuphak, C. Wasi, C. Nuchprayoon, and N. A. T. S. i. T. B. D. Working Group for.** 2009. One-year experience of nucleic acid technology testing for human immunodeficiency virus Type 1, hepatitis C virus, and hepatitis B virus in Thai blood donations. Transfusion **49:**1126-35.
234. **Posada, D., and K. A. Crandall.** 1998. MODELTEST: testing the model of DNA substitution. Bioinformatics **14:**817-8.
235. **Pourkarim, M. R., P. Lemey, S. Amini-Bavil-Olyaee, P. Maes, and M. Van Ranst.** 2010. Novel hepatitis B virus subgenotype A6 in African-Belgian patients. J Clin Virol **47:**93-6.
236. **Poussin, K., H. Dienes, H. Sirma, S. Urban, M. Beaugrand, D. Franco, P. Schirmacher, C. Brechot, and P. Paterlini Brechot.** 1999. Expression of mutated hepatitis B virus X genes in human hepatocellular carcinomas. Int J Cancer **80:**497-505.
237. **Pybus, O. G., P. V. Markov, A. Wu, and A. J. Tatem.** 2007. Investigating the endemic transmission of the hepatitis C virus. Int J Parasitol **37:**839-49.
238. **Quarleri, J., F. Moretti, M. B. Bouzas, N. Laufer, M. G. Carrillo, S. F. Giuliano, H. Perez, P. Cahn, and H. Salomon.** 2007. Hepatitis B virus genotype distribution and its lamivudine-resistant mutants in HIV-coinfected patients with chronic and occult hepatitis B. AIDS Res Hum Retroviruses **23:**525-31.
239. **Quintero, A., D. Martinez, B. Alarcon De Noya, A. Costagliola, L. Urbina, N. Gonzalez, F. Liprandi, D. Castro De Guerra, and F. H. Pujol.** 2002. Molecular epidemiology of hepatitis B virus in Afro-Venezuelan populations. Arch Virol **147:**1829-36.
240. **Quintero, A., N. Uzcategui, C. L. Loureiro, L. Villegas, X. Illarramendi, M. E. Guevara, J. E. Ludert, L. Blitz, F. Liprandi, and F. H. Pujol.** 2001. Hepatitis delta virus genotypes I and III circulate associated with hepatitis B virus genotype F In Venezuela. J Med Virol **64:**356-9.
241. **Radjef, N., E. Gordien, V. Ivaniushina, E. Gault, P. Anais, T. Drugan, J. C. Trinchet, D. Roulot, M. Tamby, M. C. Milinkovitch, and P. Deny.** 2004. Molecular phylogenetic analyses indicate a wide and ancient radiation of African hepatitis delta virus, suggesting a deltavirus genus of at least seven major clades. J Virol **78:**2537-44.
242. **Raimondo, G., J. P. Allain, M. R. Brunetto, M. A. Buendia, D. S. Chen, M. Colombo, A. Craxi, F. Donato, C. Ferrari, G. B. Gaeta, W. H. Gerlich, M. Levrero, S. Locarnini, T. Michalak, M. U. Mondelli, J. M. Pawlotsky, T. Pollicino, D. Prati, M. Puoti, D. Samuel, D. Shouval, A. Smedile, G. Squadrito, C. Trepo, E. Villa, H. Will, A. R. Zanetti, and F. Zoulim.** 2008. Statements from the Taormina expert meeting on occult hepatitis B virus infection. J Hepatol **49:**652-7.
243. **Rehermann, B., and M. Nascimbeni.** 2005. Immunology of hepatitis B virus and hepatitis C virus infection. Nat Rev Immunol **5:**215-29.
244. **Robinson, W. S., D. A. Clayton, and R. L. Greenman.** 1974. DNA of a human hepatitis B virus candidate. J Virol **14:**384-91.

245. **Robinson, W. S., and R. L. Greenman.** 1974. DNA polymerase in the core of the human hepatitis B virus candidate. J Virol **13:**1231-6.
246. **Roingeard, P., A. Diouf, J. L. Sankale, C. Boye, S. Mboup, F. Diadhiou, and M. Essex.** 1993. Perinatal transmission of hepatitis B virus in Senegal, west Africa. Viral Immunol **6:**65-73.
247. **Rosenblum, L., W. Darrow, J. Witte, J. Cohen, J. French, P. S. Gill, J. Potterat, K. Sikes, R. Reich, and S. Hadler.** 1992. Sexual practices in the transmission of hepatitis B virus and prevalence of hepatitis delta virus infection in female prostitutes in the United States. JAMA **267:**2477-81.
248. **Sa-Nguanmoo, P., P. Tangkijvanich, N. Thawornsuk, P. Vichaiwattana, K. Prianantathavorn, A. Theamboonlers, Y. Tanaka, and Y. Poovorawan.** 2010. Molecular epidemiological study of hepatitis B virus among migrant workers from Cambodia, Laos, and Myanmar to Thailand. J Med Virol **82:**1341-9.
249. **Sa-Nguanmoo, P., C. Thongmee, P. Ratanakorn, R. Pattanarangsan, R. Boonyarittichaikij, S. Chodapisitkul, A. Theamboonlers, P. Tangkijvanich, and Y. Poovorawan.** 2008. Prevalence, whole genome characterization and phylogenetic analysis of hepatitis B virus in captive orangutan and gibbon. J Med Primatol.
250. **Sakamoto, T., Y. Tanaka, E. Orito, J. Co, J. Clavio, F. Sugauchi, K. Ito, A. Ozasa, A. Quino, R. Ueda, J. Sollano, and M. Mizokami.** 2006. Novel subtypes (subgenotypes) of hepatitis B virus genotypes B and C among chronic liver disease patients in the Philippines. J Gen Virol **87:**1873-82.
251. **Satake, M., R. Taira, H. Yugi, S. Hino, K. Kanemitsu, H. Ikeda, and K. Tadokoro.** 2007. Infectivity of blood components with low hepatitis B virus DNA levels identified in a lookback program. Transfusion **47:**1197-205.
252. **Schaefer, S.** 2007. Hepatitis B virus taxonomy and hepatitis B virus genotypes. World J Gastroenterol **13:**14-21.
253. **Schaefer, S., and W. H. Gerlich.** 1995. In vitro transformation by hepatitis B virus DNA. Intervirology **38:**143-54.
254. **Schieck, A., A. Schulze, C. Gahler, T. Muller, U. Haberkorn, A. Alexandrov, S. Urban, and W. Mier.** 2013. Hepatitis B virus hepatotropism is mediated by specific receptor recognition in the liver and not restricted to susceptible hosts. Hepatology.
255. **Schultz, U., E. Grgacic, and M. Nassal.** 2004. Duck hepatitis B virus: an invaluable model system for HBV infection. Adv Virus Res **63:**1-70.
256. **Schulze, A., K. Mills, T. S. Weiss, and S. Urban.** 2012. Hepatocyte polarization is essential for the productive entry of the hepatitis B virus. Hepatology **55:**373-83.
257. **Schulze, A., A. Schieck, Y. Ni, W. Mier, and S. Urban.** 2010. Fine mapping of pre-S sequence requirements for hepatitis B virus large envelope protein-mediated receptor interaction. J Virol **84:**1989-2000.
258. **Scott, R. M., R. Snitbhan, W. H. Bancroft, H. J. Alter, and M. Tingpalapong.** 1980. Experimental transmission of hepatitis B virus by semen and saliva. J Infect Dis **142:**67-71.
259. **Seeger, C., D. Ganem, and H. E. Varmus.** 1986. Biochemical and genetic evidence for the hepatitis B virus replication strategy. Science **232:**477-84.
260. **Seeger, C., F. Zoulim, and W. Mason.** 2007. Hepadnaviruses. *In* D. M. Knipe, P. M. Howley, D. E. Griffin, R. A. Lamb, M. A. Martin, B. Roizman, and S. E. Straus (ed.), Fields Virology, 5th ed. ed. Lippincott, Williams & Wilkins, Philadelphia, PA.
261. **Sheldon, J., B. Rodes, F. Zoulim, A. Bartholomeusz, and V. Soriano.** 2006. Mutations affecting the replication capacity of the hepatitis B virus. J Viral Hepat **13:**427-34.

262. **Shih, H. H., C. Shih, H. W. Wang, C. W. Su, I. J. Sheen, and J. C. Wu.** 2010. Pro-205 of large hepatitis delta antigen and Pro-62 of major hepatitis B surface antigen influence the assembly of different genotypes of hepatitis D virus. J Gen Virol **91:**1004-12.
263. **Shimoda, T., T. Shikata, T. Karasawa, S. Tsukagoshi, M. Yoshimura, and I. Sakurai.** 1981. Light microscopic localization of hepatitis B virus antigens in the human pancreas. Possibility of multiplication of hepatitis B virus in the human pancreas. Gastroenterology **81:**998-1005.
264. **Silverman, A. L., L. H. Qu, J. Blow, I. M. Zitron, S. C. Gordon, and E. D. Walker.** 2001. Assessment of hepatitis B virus DNA and hepatitis C virus RNA in the common bedbug (Cimex lectularius L.) and kissing bug (Rodnius prolixus). Am J Gastroenterol **96:**2194-8.
265. **Simmonds, P., and S. Midgley.** 2005. Recombination in the genesis and evolution of hepatitis B virus genotypes. J Virol **79:**15467-76.
266. **Simonsen, L., A. Kane, J. Lloyd, M. Zaffran, and M. Kane.** 1999. Unsafe injections in the developing world and transmission of bloodborne pathogens: a review. Bull World Health Organ **77:**789-800.
267. **Sitnik, R., H. Sette, Jr., R. A. Santana, L. C. Menezes, C. H. Graca, G. T. Dastoli, S. Silbert, and J. R. Pinho.** 2007. Hepatitis B virus genotype E detected in Brazil in an African patient who is a frequent traveler. Braz J Med Biol Res **40:**1689-92.
268. **Smith, J. M.** 1992. Analyzing the mosaic structure of genes. J Mol Evol **34:**126-9.
269. **Standring, D. N., J. H. Ou, F. R. Masiarz, and W. J. Rutter.** 1988. A signal peptide encoded within the precore region of hepatitis B virus directs the secretion of a heterogeneous population of e antigens in Xenopus oocytes. Proc Natl Acad Sci U S A **85:**8405-9.
270. **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.
271. **Su, F., and R. J. Schneider.** 1996. Hepatitis B virus HBx protein activates transcription factor NF-kappaB by acting on multiple cytoplasmic inhibitors of rel-related proteins. J Virol **70:**4558-66.
272. **Sugauchi, F., H. Kumada, S. A. Acharya, S. M. Shrestha, M. T. Gamutan, M. Khan, R. G. Gish, Y. Tanaka, T. Kato, E. Orito, R. Ueda, Y. Miyakawa, and M. Mizokami.** 2004. Epidemiological and sequence differences between two subtypes (Ae and Aa) of hepatitis B virus genotype A. J Gen Virol **85:**811-20.
273. **Sugauchi, F., M. Mizokami, E. Orito, T. Ohno, H. Kato, S. Suzuki, Y. Kimura, R. Ueda, L. A. Butterworth, and W. G. Cooksley.** 2001. A novel variant genotype C of hepatitis B virus identified in isolates from Australian Aborigines: complete genome sequence and phylogenetic relatedness. J Gen Virol **82:**883-92.
274. **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.
275. **Sugauchi, F., E. Orito, H. Kato, S. Suzuki, S. Kawakita, Y. Sakamoto, K. Fukushima, T. Akiba, N. Yoshihara, R. Ueda, and M. Mizokami.** 2003. Genotype, serotype, and phylogenetic characterization of the complete genome sequence of hepatitis B virus isolates from Malawian chronic carriers of the virus. J Med Virol **69:**33-40.
276. **Summers, J., J. M. Smolec, and R. Snyder.** 1978. A virus similar to human hepatitis B virus associated with hepatitis and hepatoma in woodchucks. Proc Natl Acad Sci U S A **75:**4533-7.

277. **Sureau, C., B. Guerra, and R. E. Lanford.** 1993. Role of the large hepatitis B virus envelope protein in infectivity of the hepatitis delta virion. J Virol **67:**366-72.
278. **Suzuki, S., F. Sugauchi, E. Orito, H. Kato, S. Usuda, L. Siransy, I. Arita, Y. Sakamoto, N. Yoshihara, A. El-Gohary, R. Ueda, and M. Mizokami.** 2003. Distribution of hepatitis B virus (HBV) genotypes among HBV carriers in the Cote d'Ivoire: complete genome sequence and phylogenetic relatedness of HBV genotype E. J Med Virol **69:**459-65.
279. **Szmaragd, C., and F. Balloux.** 2007. The population genomics of hepatitis B virus. Mol Ecol **16:**4747-58.
280. **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.
281. **Tallo, T., H. Norder, V. Tefanova, T. Krispin, L. Priimagi, S. Mukomolov, M. Mikhailov, and L. O. Magnius.** 2004. Hepatitis B virus genotype D strains from Estonia share sequence similarity with strains from Siberia and may specify ayw4. J Med Virol **74:**221-7.
282. **Tamura, K., J. Dudley, M. Nei, and S. Kumar.** 2007. MEGA4: Molecular Evolutionary Genetics Analysis (MEGA) software version 4.0. Mol Biol Evol **24:**1596-9.
283. **Tamura, K., D. Peterson, N. Peterson, G. Stecher, M. Nei, and S. Kumar.** 2011. MEGA5: molecular evolutionary genetics analysis using maximum likelihood, evolutionary distance, and maximum parsimony methods. Mol Biol Evol **28:**2731-9.
284. **Tatematsu, K., Y. Tanaka, F. Kurbanov, F. Sugauchi, S. Mano, T. Maeshiro, T. Nakayoshi, M. Wakuta, Y. Miyakawa, and M. Mizokami.** 2009. A Genetic Variant of Hepatitis B Virus Divergent from Known Human and Ape Genotypes Isolated from a Japanese Patient and Provisionally Assigned to New Genotype J. J Virol.
285. **Taylor, J., P. Farci, and P. R.H.** 2007. Hepatitis D (Delta) virus, p. 3031-3046. *In* D. M. Knipe, P. M. Howley, D. E. Griffin, R. A. Lamb, M. A. Martin, B. Roizman, and S. E. Straus (ed.), Fields Virology, 5th ed. ed. Lippincott, Williams & Wilkins, Philadelphia, PA.
286. **Taylor, J. M.** 2006. Hepatitis delta virus. Virology **344:**71-6.
287. **Thedja, M. D., D. H. Muljono, N. Nurainy, C. H. Sukowati, J. Verhoef, and S. Marzuki.** 2011. Ethnogeographical structure of hepatitis B virus genotype distribution in Indonesia and discovery of a new subgenotype, B9. Arch Virol **156:**855-68.
288. **Thedja, M. D., M. Roni, A. R. Harahap, N. C. Siregar, S. I. Ie, and D. H. Muljono.** 2010. Occult hepatitis B in blood donors in Indonesia: altered antigenicity of the hepatitis B virus surface protein. Hepatol Int **4:**608-14.
289. **Thio, C. L., E. C. Seaberg, R. Skolasky, Jr., J. Phair, B. Visscher, A. Munoz, and D. L. Thomas.** 2002. HIV-1, hepatitis B virus, and risk of liver-related mortality in the Multicenter Cohort Study (MACS). Lancet **360:**1921-6.
290. **Thompson, J. D., D. G. Higgins, and T. J. Gibson.** 1994. CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. Nucleic Acids Res **22:**4673-80.
291. **Toh, H., H. Hayashida, and T. Miyata.** 1983. Sequence homology between retroviral reverse transcriptase and putative polymerases of hepatitis B virus and cauliflower mosaic virus. Nature **305:**827-9.

292. **Torbenson, M., and D. L. Thomas.** 2002. Occult hepatitis B. Lancet Infect Dis **2:**479-86.
293. **Toro, C., V. Jimenez, C. Rodriguez, J. Del Romero, B. Rodes, A. Holguin, P. Alvarez, M. Garcia-Campello, C. Gomez-Hernando, A. Guelar, J. Sheldon, C. de Mendoza, A. Simon, and V. Soriano.** 2006. Molecular and epidemiological characteristics of blood-borne virus infections among recent immigrants in Spain. J Med Virol **78:**1599-608.
294. **Tran, T. T., T. N. Trinh, and K. Abe.** 2008. New complex recombinant genotype of hepatitis B virus identified in Vietnam. J Virol **82:**5657-63.
295. **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.
296. **Van Brussel, M., M. Salemi, H. F. Liu, P. Goubau, J. Desmyter, and A. M. Vandamme.** 1999. The discovery of two new divergent STLVs has implications for the evolution and epidemiology of HTLVs. Rev Med Virol **9:**155-70.
297. **van Steenbergen, J. E., H. G. Niesters, E. L. Op de Coul, G. J. van Doornum, A. D. Osterhaus, A. Leentvaar-Kuijpers, R. A. Coutinho, and J. A. van den Hoek.** 2002. Molecular epidemiology of hepatitis B virus in Amsterdam 1992-1997. J Med Virol **66:**159-65.
298. **Venegas, M., G. Munoz, C. Hurtado, L. Alvarez, M. Velasco, R. A. Villanueva, and J. Brahm.** 2008. Prevalence of hepatitis B virus genotypes in chronic carriers in Santiago, Chile. Arch Virol **153:**2129-32.
299. **Venook, A. P., C. Papandreou, J. Furuse, and L. L. de Guevara.** 2010. The incidence and epidemiology of hepatocellular carcinoma: a global and regional perspective. Oncologist **15 Suppl 4:**5-13.
300. **Vieth, S., C. Manegold, C. Drosten, T. Nippraschk, and S. Gunther.** 2002. Sequence and phylogenetic analysis of hepatitis B virus genotype G isolated in Germany. Virus Genes **24:**153-6.
301. **Vray, M., J. M. Debonne, J. M. Sire, N. Tran, B. Chevalier, J. C. Plantier, F. Fall, G. Vernet, F. Simon, and P. S. Mb.** 2006. Molecular epidemiology of hepatitis B virus in Dakar, Senegal. J Med Virol **78:**329-34.
302. **Wang, J. T., C. Z. Lee, P. J. Chen, T. H. Wang, and D. S. Chen.** 2002. Transfusion-transmitted HBV infection in an endemic area: the necessity of more sensitive screening for HBV carriers. Transfusion **42:**1592-7.
303. **Wang, Z., Z. Liu, G. Zeng, S. Wen, Y. Qi, S. Ma, N. V. Naoumov, and J. Hou.** 2005. A new intertype recombinant between genotypes C and D of hepatitis B virus identified in China. J Gen Virol **86:**985-90.
304. **Watanabe, T., E. M. Sorensen, A. Naito, M. Schott, S. Kim, and P. Ahlquist.** 2007. Involvement of host cellular multivesicular body functions in hepatitis B virus budding. Proc Natl Acad Sci U S A **104:**10205-10.
305. **Wei, Y., C. Neuveut, P. Tiollais, and M. A. Buendia.** 2010. Molecular biology of the hepatitis B virus and role of the X gene. Pathol Biol (Paris) **58:**267-72.
306. **WHO.** 2002. Hepatitis B (WHO/CDS/CSR/LYO/2002.2).
307. **WHO.** 2004. Hepatitis B vaccines. WHO position paper. Weekly Epidemiol Rec **28:**253-264.
308. **WHO.** 2009. Hepatitis B vaccines. WHO position paper. Weekly Epidemiol Rec **40:**405-420.
309. **WHO.** 2000. Hepatitis B. WHO fact sheet No 204.
310. **WHO.** 2001. Hepatitis Delta (WHO/CDS/CSR/NCS/2001.1).
311. **Wu, J. C., C. M. Chen, I. J. Sheen, S. D. Lee, H. M. Tzeng, and K. B. Choo.** 1995. Evidence of transmission of hepatitis D virus to spouses from sequence analysis of the viral genome. Hepatology **22:**1656-60.

312. **Wu, J. C., T. Y. Chiang, and I. J. Sheen.** 1998. Characterization and phylogenetic analysis of a novel hepatitis D virus strain discovered by restriction fragment length polymorphism analysis. J Gen Virol **79 (Pt 5):**1105-13.
313. **Xu, Z., T. S. Yen, L. Wu, C. R. Madden, W. Tan, B. L. Slagle, and J. H. Ou.** 2002. Enhancement of hepatitis B virus replication by its X protein in transgenic mice. J Virol **76:**2579-84.
314. **Yang, J., K. Xing, R. Deng, J. Wang, and X. Wang.** 2006. Identification of Hepatitis B virus putative intergenotype recombinants by using fragment typing. J Gen Virol **87:**2203-15.
315. **Yeh, C. T.** 2010. Development of HBV S gene mutants in chronic hepatitis B patients receiving nucleotide/nucleoside analogue therapy. Antivir Ther **15:**471-5.
316. **Yu, H., Q. Yuan, S. X. Ge, H. Y. Wang, Y. L. Zhang, Q. R. Chen, J. Zhang, P. J. Chen, and N. S. Xia.** 2010. Molecular and phylogenetic analyses suggest an additional hepatitis B virus genotype "I". PLoS One **5:**e9297.
317. **Yu, M. C., J. M. Yuan, S. Govindarajan, and R. K. Ross.** 2000. Epidemiology of hepatocellular carcinoma. Can J Gastroenterol **14:**703-9.
318. **Yuan, Q., S. H. Ou, C. R. Chen, S. X. Ge, B. Pei, Q. R. Chen, Q. Yan, Y. C. Lin, H. Y. Ni, C. H. Huang, A. E. Yeo, J. W. Shih, J. Zhang, and N. S. Xia.** 2010. Molecular characteristics of occult hepatitis B virus from blood donors in southeast China. J Clin Microbiol **48:**357-62.
319. **Yuen, M. F., E. Sablon, Y. Tanaka, T. Kato, M. Mizokami, J. Doutreloigne, H. J. Yuan, D. K. Wong, S. M. Sum, and C. L. Lai.** 2004. Epidemiological study of hepatitis B virus genotypes, core promoter and precore mutations of chronic hepatitis B infection in Hong Kong. J Hepatol **41:**119-25.
320. **Zaaijer, H. L., S. Bouter, and H. J. Boot.** 2008. Substitution rate of the hepatitis B virus surface gene. J Viral Hepat **15:**239-45.
321. **Zeng, G. B., S. J. Wen, Z. H. Wang, L. Yan, J. Sun, and J. L. Hou.** 2004. A novel hepatitis B virus genotyping system by using restriction fragment length polymorphism patterns of S gene amplicons. World J Gastroenterol **10:**3132-6.
322. **Zhang, M., A. K. Schultz, C. Calef, C. Kuiken, T. Leitner, B. Korber, B. Morgenstern, and M. Stanke.** 2006. jpHMM at GOBICS: a web server to detect genomic recombinations in HIV-1. Nucleic Acids Res **34:**W463-5.
323. **Zhang, Q., and G. Cao.** 2011. Genotypes, mutations, and viral load of hepatitis B virus and the risk of hepatocellular carcinoma: HBV properties and hepatocarcinogenesis. Hepat Mon **11:**86-91.
324. **Zhang, Y. Y., E. Tsega, and B. G. Hansson.** 1996. Phylogenetic analysis of hepatitis D viruses indicating a new genotype I subgroup among African isolates. J Clin Microbiol **34:**3023-30.
325. **Zhang, Y. Y., B. H. Zhang, D. Theele, S. Litwin, E. Toll, and J. Summers.** 2003. Single-cell analysis of covalently closed circular DNA copy numbers in a hepadnavirus-infected liver. Proc Natl Acad Sci U S A **100:**12372-7.
326. **Zhang, Z., U. Protzer, Z. Hu, J. Jacob, and T. J. Liang.** 2004. Inhibition of cellular proteasome activities enhances hepadnavirus replication in an HBX-dependent manner. J Virol **78:**4566-72.
327. **Zhou, Y., and E. C. Holmes.** 2007. Bayesian estimates of the evolutionary rate and age of hepatitis B virus. J Mol Evol **65:**197-205.
328. **Zoulim, F., J. Saputelli, and C. Seeger.** 1994. Woodchuck hepatitis virus X protein is required for viral infection in vivo. J Virol **68:**2026-30.
329. **Zuckerman, J. N.** 2005. Prevention, p. 370-377. *In* H. C. Thomas, S. M. Lemon, and A. J. Zuckerman (ed.), Viral Hepatitis. Blackwell Publishing Ltd, Oxford, UK.

330. **Zuckerman, J. N., and A. J. Zuckerman.** 2003. Mutations of the surface protein of hepatitis B virus. Antiviral Res **60:**75-8.

Annex

1 Conference Participations

2011 15th Saar-Lor-Lux-Meeting on Virus Research. Remich, Luxembourg.

2011 PhD Day Life Sciences. Luxembourg, Luxembourg.

2011 21st Annual Meeting of the Society for Virology (GfV). Freiburg, Germany.

2011 Euroscience Open Forum. Turin, Italy.

2010 14th Saar-Lor-Lux-Meeting on Virus Research. Homburg, Germany.

2010 PhD Day Life Sciences. Luxembourg, Luxembourg.

2010 Simian viruses and emerging diseases in humans. Virology, epidemiology and medical history. International Symposium. Paris, France. Invited Presentation.

2009 2nd annual Bioinformatics Symposium. LuciLinX 2009. Luxembourg, Luxembourg.

2009 50th Doctoral Student at the Institute of Immunology. Scientific Symposium. Luxembourg, Luxembourg.

2009 29iéme Journée Nationale de Biologie Clinique. Luxembourg, Luxembourg.

2009 13th Saar-Lor-Lux-Meeting on Virus Research. Nancy, France.

2009 International Meeting - The Molecular Biology of Hepatitis B Viruses. Tours, France.

2008 12th Saar-Lor-Lux-Meeting on Virus Research. Remich, Luxembourg.

2007 International Meeting - The Molecular Biology of Hepatitis B Viruses. Rom, Italy.

2007 11th Saar-Lor-Lux-Meeting on Virus Research. Homburg, Germany.

2 Publications

2013 **Andernach IE**, Hunewald O, Muller CP. Bayesian inference of the evolution of HBV/E in sub-Saharan Africa. PLoS ONE. Submitted.

2013 **Andernach IE**, Leiß L, Tarnagda Z, Tahita CM, Otegbayo J, Forbi JC, Omilabu S, Komas NC, Gouandjika I, Mbah PO, Muller CP. Hepatitis Delta Virus (HDV) in sub-Saharan Africa. Manuscript in preparation.

2013 Jutavijittum P, **Andernach IE**, Samountry B, Yousukh A, Thammavong T, Muller CP. Serological characterization of HBV in Laos. Vox Sanguinis. Accepted.

2012 **Andernach IE**, Jutavijittum P, Samountry B, Yousukh A, Thammavong T, Hübschen JM, Muller CP. A cauldron of mixed infections and recombinations of hepatitis B virus in Laos. PLoS ONE, 7(2): e30245.

2009 **Andernach IE**, Nolte C, Pape JW, Muller CP. Slave Trade and Hepatitis B Virus Genotypes and Subgenotypes in Haiti and Africa. Emerg Infect Dis. 15(8):1222-28.

2009 **Andernach IE**, Hübschen JM, Muller CP. Hepatitis B virus: the genotype E puzzle. Rev Med Virol. 19(4):231-40.

2008 Hübschen JM, **Andernach IE**, Muller CP. Hepatitis B virus genotype E variability in Africa. J Clin Virol. 43:376–380.

2006 Schönert S, Seitz S, Krafft H, Feuerbaum EA, **Andernach I**, Witz G, Dahl MK. Maltose and maltodextrin utilization by Bacillus subtilis. J Bacteriol. 188(11):3911-22.

3 Acknowledgements

I wish to thank Prof. Dr. Claude P. Muller and Prof. Dr. Nikolaus Müller-Lantzsch for giving me the opportunity to conduct my dissertation under their supervision. I am grateful for Prof. Mullers continuous scientific guidance and expertise throughout all stages of my PhD study.

I furthermore wish to thank Dr. Judith Hübschen, who was always available to discuss ideas and to give invaluable expertise and advice, and Dr. Christophe Olinger for scientific advice and technical expertise.

I am very grateful for the financial support of the "Bourse de Formation Recherche" and "Aides à la Formation Recherche" of the Ministry for Higher Education and Research, the Fonds National de la Recherche. I also thank the Recherche Public de la Santé (CRP-Santé) for the support and financial contributions.

I would especially like to thank the numerous international collaborators that were involved in sample collection, analyses and publications, especially Assoc. Prof. Dr. Prapan Jutavijittum, Prof. Dr. Jean W. Pape, Claudine A. Nolte and Dr. Patrick O. Mbah. Many thanks also to all visiting scientists and students. It was a pleasure working with all of you!

Also, I would like to thank the great team at the Institute of Immunology, especially Anna for being the best office mate ever and for invaluable discussions, support and friendship. Special thanks also go to Ulla and Carole for their administrative and organisational support, Aurélie, Emilie, and Sébastien for their excellent help and technical advice and Oliver for the great support and the assistance with computer

problems. A special thank you also to all present and former colleagues at the Institute: Andrea, Anja, Anna, Antony, Aurélie, Axel, Brian, Carlos, Carole, Chantal, Christophe, Claire, Denis, Dominique, Emilie, Fabienne, Fred, Jacques, Joana, Jon, Joséphine, Josiane, Judith, Julia, Konstantin, Laetitia, Lei, Linda, Lukas, Marianna, Mariette, Mario, Monique, Nadia, Nancy, Nathalie, Nick, Nina, Oliver, Patrick, Regina, Sandra, Sara, Sébastien, Simone A., Simone R., Slavena, Sophie F., Sophie K., Sophie M., Stefanie, Steffi, Stéphanie, Tom, Ulla, Victor and Vitor.

I also wish to thank all my friends who made my life enjoyable during and after work.

Most of all I wish to thank my parents and my brother for their love, encouragement and unconditional support.