High prevalence of hepatitis B virus genotype E in Northern Madagascar indicates a West-African lineage
Tatiana Dupinay, Karin Restorp, Peter Leutscher, Dominique Rousset, Isabelle Chemin, Rene Migliani, Lars O Magnius, Helene Norder

To cite this version:

HAL Id: hal-00599781
https://hal.archives-ouvertes.fr/hal-00599781
Submitted on 11 Jun 2011

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L’archive ouverte pluridisciplinaire HAL, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d’enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.
High prevalence of hepatitis B virus genotype E in Northern Madagascar indicates a West-African lineage
High prevalence of hepatitis B virus genotype E in Northern Madagascar indicates a West-African lineage

Tatiana Dupinay¹²³, Karin Restorp⁴, Peter Leutscher⁵, Dominique Rousset⁷, Isabelle Chemin¹, Rene Migliani⁵, Lars Magnius⁴, Heléne Norder¹*.  

¹ INSERM U871, 151 Cours Albert Thomas, 69003 Lyon, France;  
² Université Lyon 1, IFR62 Lyon Est, 69008 Lyon, France;  
³ Ecole pratique des hautes études, 75007 Paris France  
⁴ Virological Department, Swedish Institute of Infectious Disease Control, Sweden  
⁵ Institute Pasteur de Madagascar, Antananarvio, Madagascar  
⁶ Department of Infectious Diseases, Aarhus University Hospital, Skejby, DK-8200 Aarhus N, Denmark  
⁷ Institute de Medicine Tropicale du Service de Sante des Armées, Marseille, France.

* Correspondence to: Helene Norder  
Department of Virology  
Swedish Institute for Infectious Disease Control, SE-171 82 Solna, Sweden  
e-mail: helene.norder@smi.se

Word Count
Abstract: 249 words
Main text: 3820 words
4 Tables and 1 Figure
Running title: HBV genotype E in Madagascar
Abstract

The prevalence of hepatitis B virus (HBV) markers was investigated in 566 inhabitants aged 15 to 55 years from a sugar cane region, Sirama, and from a village, Mataipako, in Northern Madagascar. Serological markers of past or present infection were significantly higher in Sirama, 74 % versus 55 %. There was no difference in the prevalence of chronic HBsAg carriers, 8.7 versus 8.5 % between the two regions. Sequencing the S gene in 45 strains revealed a predominance of genotype E, in 53 %, followed by subgenotype A1 in 22 %, and genotype D in 18 %. Phylogenetic analyses of the genotype E strains showed homology with West African strains. All A1 isolates were similar to Malawi strains. Most genotype D strains were subgenotype D7 and related to strains from Somalia and Tunisia. One genotype D strain formed a branch between Pacific D4 and African D7 strains at neighbour-joining analysis. The precore stop mutant was found in 33 % of the genotype D strains, 17 % of E but not in any A1 strain. The high prevalence and low variability of genotype E strains in only two villages, indicates a rather recent introduction of this genotype into Madagascar from West Africa, possibly through migration or slave trade. The wider spread and genetic relationship of genotype D with East African and Austronesian strains indicate an earlier introduction of this genotype. Molecular epidemiology of HBV may thus be used to complement linguistic and genetic studies on past human migrations in Africa.
**Introduction**

Hepatitis B virus (HBV) is a human pathogen causing both acute and chronic hepatitis. HBV is a major public health concern and approximately two billion people worldwide have serological markers of past or present infection. More than 350 million people are chronically infected according to the World Health Organisation (WHO). In addition, WHO estimations indicate that approximately one third of all cases of cirrhosis and half of all cases of hepatocellular carcinoma can be attributed to chronic HBV infection worldwide. The prevalence of chronic infection varies throughout the world with South East Asia, China, Amazon Basin and Africa carrying the highest burden of disease. Parenteral and sexual transmissions are globally the most common modes of transmission. In endemic countries the predominant routes of transmission are perinatal transmission from mother to child during birth in Asia, and horizontal transmission between children in Africa. A global hepatitis B immunization program has been recommended by WHO since 1991.

Madagascar is a country with an ethnically diverse population divided into 22 tribes and a growth rate of 2.8 % per year (Hewitt et al., 1996; Mafrezi and Randretsa, 1985). Madagascar is classified as an area of high endemicity for HBV, and about 5% of the urban population and 30% of the rural population are chronic HBsAg carriers (Boisier et al., 1996; Migliani et al., 2000a; Migliani et al., 2000b; Morvan et al., 1994). Hepatitis B vaccination is the most effective strategy for prevention of hepatitis B. However, in Madagascar laboratory facilities are insufficient and the use of hepatitis B vaccine is limited (Migliani et al., 2000b).

HBV is a small, double-stranded DNA virus and the prototype of the hepadnavirus family (Ganem & Varmus, 1987). HBV strains are classified into eight genotypes, designated A-H (Arauz-Ruiz et al., 2002; Naumann et al., 1993; Norder et al., 1993; Okamoto et al., 1987; Stuyver et al., 2000). Each genotype has distinct geographical distribution. Genotype A is common in North-Western Europe, North America, and Africa. Genotypes B and C are predominant in East- and South East Asia, Oceania and among aboriginal populations in Australia. Genotype D is found worldwide and is predominant in the Mediterranean, Middle East, and India, whereas genotype E has so far been confined to populations in western parts of sub-Saharan Africa apart from one report from India (Norder et al., 2004; Singh et al., 2009). Genotype G has been identified in samples from several countries, and is found in patients co-infected with another HBV genotype. Genotypes F and H are found mainly in Central and South America, and are the indigenous genotypes of Amerindian populations (Arauz-Ruiz et al., 2002; Norder et al., 2004; Kumagai et al., 2007; Sanchez et al., 2007).
The major HBV genotypes apart from E and H have been divided into subgenotypes based on phylogenetic analyses and genetic divergence of complete genomes. Each subgenotype has specific geographical distribution. Six subgenotypes have been recognized for genotype A. A1 is prevalent in Africa, Asia and Indonesia (Bowyer et al., 1997; Kramvis et al., 2002; Banerjee et al., 2006; Kramvis and Kew 2007; Tanaka et al., 2004; Hannoun et al., 2005). A2 is prevalent in North West Europe (McMahon 2009; Norder et al., 2004; Sugauchi et al., 2004; Schaefer, 2007). A3 – A6 are found in West and Central Africa, (Makuwa et al., 2006; Kramvis and Kew 2007; Pourkarim et al., 2009; Kurbanov et al., 2005; Olinger et al., 2006). There are eight subgenotypes for genotype B, and seven for C. B1 and C2 are prevalent in Japan, Korea and Northern China, C2 is also found in Alaska, while B2 and C1 are prevalent in Southern China, Taiwan, and South East Asia (Norder et al., 2004; Schaefer, 2007; McMahon, 2009; Wang et al., 2009). B3 – B5 and B8 as well as C3 – C7 are prevalent in the Pacific, while B6 is found in Alaska, Northern Canada and Greenland (McMahon, 2009; Mulyanto et al., 2009; Cavinta et al., 2009; Lusida et al., 2008; Nagasaki et al., 2006; Sakamoto et al., 2006). Subgenotype B7 described from Indonesia (Nurainy et al., 2008) probably represents strains of subgenotype B3. There are seven subgenotypes reported for genotype D. D1 is highly prevalent in the Middle East and in the Mediterranean countries, but also occurs in India and East Africa and (Banerjee et al., 2006; Norder et al., 2004; Schaefer 2007). D2 is prevalent in East Europe including Russia and the Baltic States, while D3 has a more South-Eastern distribution in Russia and is also prevalent in Northern India and Pakistan (Norder et al., 2004; Tallo et al., 2008). D3 has also a global spread attributed to its spread among injecting drug users. D5 is reported from Eastern India (Banerjee et al., 2006; Chandra et al., 2009). D4 was defined based on complete genomes from Australia and Western Pacific and formed a clade with strains from North-Africa including Somalia in genetic trees based on the small S gene (Norder et al. 2004). The North-African subclade, however, is divergent sufficiently to represent a new subgenotype, D7, based on analysis of strains from Tunisia (Meldal et al., 2009). D6 reported from Papua (Lusida et al., 2008) probably represents strains of subgenotype D4. The subgenotypes of F are found in the Americas. F1 has a Northern distribution in Alaska, Mexico, and Central America and has spread to Peru and Argentina, while F2 – F4 are prevalent in South America (McMahon, 2009; Norder et al., 2004; Devesa et al., 2008).

In this study, the prevalence of HBV markers was investigated in populations in a rural area in the Northern part of Madagascar. The HBV strains from chronic carriers were characterized in order to determine the HBV genotypes and subgenotypes prevalent in this region of Madagascar.
Material and Methods

Materials

As a part of a schistosomiasis morbidity study, sera were collected from 563 inhabitants aged 15 to 55 years living in a rural area in Northern Madagascar (Leutscher et al., 2008; Ramarakoto et al., 2008). Participants were from the Mataipako village and from five villages (Ambodikatakata, Ambodimanga, Ankatoko, Antsaboahitra, and Tanambao) in the neighboring Sirama sugarcane plantation area close to the major town Ambilobe. Two-hundred-fifty-one samples (43 %) were from Mataipako and 322 from the Sirama villages (Table I). Three hundred and one (54 %) were from males and 262 (46 %) from females (Table II). Median age of the investigated inhabitants was 29 years for both regions, 29.8 for males and 28.6 years for females. No prior HBV vaccination program or other control interventions have been undertaken in the study area. Informed consent was obtained from the participants prior to enrolment in accordance with the study protocol, which was approved by the committee of ethics in Madagascar.

Serological assays

Sera were tested by enzyme immunoassays (EIA) with commercial kits for HBsAg detection (Sanofi Diagnostics Pasteur, Marnes La Coquette, France; Murex Diagnostics, Maidenhead, England; Ortho Diagnostics Systems, New Brunswick, NJ, USA; and DiaSorin, Saluggia, Italy). Samples were also tested for anti-HBc (Wellcozyme, Abbott Diagnostics, Maidenhead, England, and Anti-HBc PLUS BioRad, Hercules, CA, US). HBeAg and anti-HBe were determined with HBeAg-Ab Plus BioRad Monolisa, and anti-HBs was detected with the EIA Anti-HBs BioRad (BioRad LaboratoriesAB, Hercules, CA, US).

Extraction of HBV DNA, PCR amplification and sequencing of the S gene, the pre-core and the core promoter regions.

HBV DNA was extracted from 50 µl serum by the standard phenol-chloroform method. Briefly, 50 µl serum were incubated for 2 hours at 42°C with 250 mg/ml of proteinase K (Sigma, St Louis, MO, US), 1% sodium dodecysulphate, 2.5 mM disodium EDTA (Sigma, St Louis, MO, US), 25 mM sodium acetate and 0.25 µg/ml tRNA. The DNA was subsequently extracted with phenol and with a 1:1 mixture of phenol:chloroform. Thereafter the DNA was precipitated in absolute ethanol with 0.15 M of potassium acetate and dissolved in 40 µl of milli-Q water.

PCR amplification of the S-region was performed by using primers hep75 and 73b in the first PCR and hep3/hep33 and hep4/hep34 for nesting (Norder et al., 1992). The pre-core region was amplified
with primer promC and hep68 (Sendi et al., 2005), and nested with primers hep67 and hep68 (Arauz-Ruiz et al., 1997). For amplification of the core promoter region, primers hep64 and hep70 were used in the first round of PCR and hep64 and hep54 in the nested PCR (Sendi et al., 2005).

Purified products were used as templates in the sequencing reaction using the dideoxy-nucleotide chain-termination method with an ABI PRISM BigDye terminator cycle-sequencing reaction kit (version 3; Applied Biosystems, Foster City, CA, US). The primers used in the PCR were used as sequencing primers. All amplification products were sequenced bi-directionally.

An ABI PRISM 3100 genetic analyser (Applied Biosystems) was used for electrophoresis and data collection.

**Phylogenetic analysis**

Sequences obtained were edited manually using the SeqMan program in the LASERGENE package (DNASTAR, Inc., Madison, WI). The sequences were thereafter aligned with the corresponding region in sequences retrieved from GenBank.

Phylogenetic analysis was carried out with the PHYLIP program package version 3.53 (Felsenstein, 1993). Evolutionary distances were estimated with the DNADIST program using the F81 model. Phylogenetic trees were constructed using UPGMA and neighbour-joining (N-J) algorithm in the PHYLIP package. Genotypes and subgenotypes were determined by phylogenetic analysis of the amplified fragments of the S gene with sequence from previously genotyped and subgenotyped strains (Norder et al., 2004). The deduced amino acid sequence of the S gene region was used to determine the subtype, which was assessed from the substitutions at codons 122, 127, and 160 (Norder et al., 1992; Okamoto et al., 1987).

**Results**

**Seroprevalence of hepatitis B markers**

There was an overall high prevalence of HBV markers (66 %). It was significantly higher in the Sirama villages, 74 % versus Mataipako, 55 %, (p<0.0001; Fisher’s exact test; Table I), whereas there was no major difference between males and females (69% versus 65%, p=0.34). In all, 370 inhabitants had HBV markers, 322 (57 %) were negative for HBsAg but had anti-HBc and/or anti-HBs as markers of past infection, while 48 (8.5 %) individuals were positive for HBsAg (Table I). Anti-HBs as only serological marker was found in nine (1.6 %) inhabitants (Table I). There was a significantly higher rate of HBV markers among males in the Sirama villages compared to Mataipako village, 70% versus
54%, (p<0.0001; Fisher’s exact test). There was no correlation between age and prevalence of HBV markers (Table I).

Among the 48 HBsAg positive patients, 20 (42%) were HBeAg positive, another 26 (54%) were anti-HBe positive, while two patients were negative for both these markers (Table II). There was no significant difference in HBeAg or anti-HBe prevalence between regions or sexes.

**Genotype distribution**

Three HBV genotypes A, D, and E, were identified in 45 samples by sequencing and analyzing the S-gene region (Table III). Most strains belonged to genotype E, 24/45 (53%), followed by genotype D, 11/45 (24%) and A, 10/45 (22%). One individual was infected dually with one genotype A strain expressing ayw1 and one genotype D strain expressing ayw2. The HBV genotype could not be determined for two individuals due to low level HBV DNA.

Most of the 24 genotype E strains originated from inhabitants in two villages, 12 from Mataipako, and 10 from one village in the Sirama region (Table III). Seven of the genotype A strains were from inhabitants in two villages in the Sirama region, and three were from inhabitants in Mataipako. The genotype D strains originated from inhabitants in all villages with HBsAg positive individuals (Table III).

**Phylogenetic analysis**

Phylogenetic analysis of the small S gene revealed that all genotype A isolates belonged to subgenotype A1 and specified subtype ayw1. In the phylogenetic tree, these isolate segregated into one clade and being most similar to strains from Malawi (Fig 1a). This clade was also supported by N-J analysis (data not shown). Five strains from inhabitants in two villages in the Sirama region, Ambodimanga and Ambodikatakata, had identical S gene sequences and differed by only one nucleotide from two other strains from Mataipako (Fig 1a). Another three strains were more divergent, although they were found in the same major cluster as the other genotype A strains from Madagascar.

All but two genotype D strains could be classified into two subgenotypes, D2 (one strain) and D7 (8 strains) in the phylogenetic analyses. The D7 strains, specifying ayw2, were found in one cluster and shared similarity with strains from Somalia and Tunisia in both the UPGMA and N-J analyses (Figs 1b; 1c). These D7 isolates originated from inhabitants in three villages, two in the Sirama region (Ambodikatakata, and Tanambao), and Mataipako. The D2 strain specifying ayw3 was from a male
from Ankatoko in the Sirama region, and was similar to strains from East Europe (Figs 1b; 1c). The D2 and D7 strains formed separate clades also in the N-J analysis, although in this analysis the D1 and D3 strains were split into two and three clades, respectively (fig. 1c). The D4 and D7 strains formed the first split in the N-J tree, although this was not the case in the UPGMA tree. Six D1 and two D2 strains classified based on complete genomes, could not be classified into a subgenotype by UPGMA but belonged to their respective subgenotypes in the N-J analysis (figs 1b; 1c). On the other hand there were five D1 strains based on complete genomes that were classified as D1 in the UPGMA tree but could not be classified into a subgenotype by N-J (figs 1b; 1c). Two genotype D strains from Madagascar were not classifiable with any of the analyses. One, 1500-02, from a female from Mataipako encoding ayw4, was found between genotype D and E strains in the N-J tree (fig 1c). The other strain, 600-02, from a male from Ambodimanga specified ayw2. This strain was similar to a strain from aboriginal populations in Australia in the UPGMA tree and formed a separate branch between the Asian D4 and the African D7 strains in the N-J tree (figs 1b; 1c).

The 24 genotype E isolates were found intermixed with strains from West Africa in the phylogenetic tree (Fig. 1d). Fourteen of the Madagascar strains were found in two clades. In the major clade, 11 of the Madagascar strains were intermixed with strains from Cameroon, Nigeria, Ivory Coast, and Mali (Fig. 1d). Eight of these strains were from Mataipako, and four from two villages in the Sirama region, Tanambao and Ankatoko. In the second cluster 4 isolates from Tanambao were intermixed with strains from Congo (Fig 1d).

Pre-core and core promoter region

The pre-core region was sequenced in 36 strains. A premature stop codon related to the G1896A mutation was observed in six sequences (16.6%). This pre-core mutant was found in three of seven investigated genotype D7 strains and in three of 18 genotype E strains. It was not found in any of nine genotype A strains investigated. All strains with the pre-core stop mutant were from patients with anti-HBe. Another pre-core mutation, G1896T, converting Trp1896 to Leu1896, was found in two genotype D strains and in one genotype E strain. This variant was found in two patients with anti-HBe and in one patient lacking HBe markers.

The core promoter region was amplified and sequenced in 16 strains, three were genotype A, another three were genotype D, and 10 were genotype E strains. All genotype D and E strains expressed A1757, and the genotype A strains expressed G1757. One genotype A strain had C1766T and T1768A
mutations in the core promoter region, and one genotype E strain had a triple mutation, A1762C, G1764T and C1766G. All other strains had wild type promoter sequence.

Discussion

The prevalence of hepatitis B markers in the rural populations from the Northern region of Madagascar was high, although not as high as the prevalence described in other rural population in the centre of Madagascar; where up to 18% were HBsAg positive (Boisier et al., 1996; Migliani et al., 2000a; Morvan et al., 1994). Only 1.6% of the inhabitants from this study had anti-HBs only which strongly suggest that implementation of control programs for hepatitis B in these areas is warranted. The present findings on the epidemiological features of persons infected with hepatitis B such as age, sex, and origin showed that the prevalence of infection appears to be higher among males from rural villages.

In this study, there was a high frequency of HBeAg positive HBsAg carriers (42%), mainly among those infected with genotypes A1 and E. This prevalence of HBeAg was higher than that from other parts of East and Central Africa as Zimbabwe, Cameroon, Ethiopia, Benin and the Central African Republic (Tswana et al., 1996; Rapicetta et al., 1991; Abebe et al., 2003; Fujiwara et al., 2005; Pawlotsky et al., 1995; Abiodun et al., 1994). In addition, there was a low prevalence of strains with mutations in the pre-core and core promoter despite the relatively high presence of A1 and D strains.

In Mediterranean countries, genotype D has been shown to present in high frequency in negative chronic HBV infection without HBeAg, associated with HBV mutants in the pre-core region (Brunetto et al., 1993). The most common precore mutation, G1896A, which creates a stop codon, has been associated previously with the subgenotypes found in this study (Chu et al., 2002; Kramvis et al., 1997; Lok et al., 1994; Olinger et al., 2006). This mutant was only found in 38% of the genotype D strains and in none of the A1 strains. There were also only two strains, one genotype A and one genotype E, with core promoter mutations. The high HBeAg prevalence and low frequency of precore and core promoter may be due to the prevalent HBV genotypes if genotypes A1 and E have prolonged replicative phase with HBeAg expression, as is described for genotypes B and C in Asia (Kao, 2002; Orito et al., 2001). It is not known if chronic carriers infected with A1 and E have a higher prevalence of HBeAg positivity than those infected with genotype D also in other parts of Africa, or if it is specific for this studied population.

The Malagasy population has a widespread origin from Asiatic to sub-Saharan African descendants and a mixture of the two (Hurles et al., 2005; Singer et al., 1957). Although both African and
Southeast Asian populations have contributed to Madagascar’s gene pool, those of Asian origin are most predominant in the central highlands, while the coastal populations often are of African origin (Chow et al., 2005; Regueiro et al., 2008; Hurles et al., 2005).

The origin of the population may also be reflected by the prevalent HBV genotypes. The subgenotype A1 strains were similar and mainly found in two villages. This subgenotype dominates in all East African countries including South Africa, Malawi, Tanzania, Uganda, and Somalia (Kramvis & Kew, 2007). The isolates from Madagascar were similar to Malawi isolates, suggesting that this subgenotype was introduced into Madagascar from East Africa. Geographically Malawi is close to the North Western part of Madagascar and historically, in the nineteenth century, the trade route from Malawi to the western coast of Madagascar became exploited increasingly for slave trading.

The genotype E strains in this study showed high sequence similarity to West African strains, where this genotype is confined. Previous studies have shown a low genetic diversity of genotype E strains isolated from a large geographical area in West Africa (Kramvis et al., 2005; Quintero et al., 2002). Several studies have proposed that this low genetic divergence together with few findings of genotype E in the Americas indicates a recent introduction of genotype E into the human population (Mulders et al., 2004; Odemuyiwa et al., 2001; Suzuki et al., 2003). The genotype E strains in this study were mainly from inhabitants in two villages, which may indicate a recent introduction, and that sufficient time has not yet passed for its subsequent spread into the rest of the country. The inhabitants in the two villages were working in the rice fields, and they may be slave descendant with origin from mainland Africa. West African origin of some Malagasy populations is supported by genetic studies on Y-chromosomal polymorphism and mitochondrial sequence diversities and on β-globulin haplotypes (Hurles et al., 2005; Hewitt et al., 1996). In addition, hepatitis C virus genotype 2 strains, common in West Africa, have been isolated from Malagasy populations, further supporting previous human migrations from West Africa to Madagascar (Markov et al., 2009).

Interestingly, 24% of the strains in this study were genotype D, and found in all five villages with HBsAg positive individuals. The majority were subgenotype D7, which are also found in Somalia and Tunisia. This is consistent with the subdivision of the previous classified D4 strains into two subgenotypes, D4 represented by strains from the Australia and Western Pacific, and D7 with North African strains (Meldal et al. 2009; Schaefer et al., 2009). The D4 and D7 strains formed the first split in the genotype D phylogenetic tree, when N-J algorithm was used, as they do when complete genomes are analysed (Meldal et al., 2009; Norder et al., 2004; Tallo et al., 2008). N-J analyses of
genotype D –genes thus seems better to mirror the tree obtained from complete genomes than
UPGMA. The split of genotype D strains into two main branches, one with D4 and D7 and one with
the other D subgenotypes, with D5 as first split, might support that genotype D has followed two waves
of two human migrations out of Africa. The D4 subgenotype being linked to the early settlement of
Papua-Australia-Melanesia followed by a later wave of settlement reaching India supported by India,
being populated before Europe, and D5 being the first split in this branch.

There were in this study five different D7 strains, all showing close relationship with strains from
Somalia. It has been proposed that there has been a migration wave into Madagascar from central East
Africa based on human genetic and linguistic studies (Dahl, 1988; Hewitt et al., 1996; Hurles et al.,
2005). Interestingly, there was one genotype D strain similar showing a distant genetic relationship to
a strain from Australia, and formed a separate branch between the African D7 and the Asian D4
strains, suggesting also Asian import of genotype D into Madagascar. This is in agreement with

genetic studies showing a South East Asian and Pacific origin of some Malagasy populations (Hurles
et al., 2005; Migot et al., 1995; Soodyall et al., 1995). The dissemination of genotype D in the
majority of the villages and its higher genetic variability compared to genotype A and E, indicate an
earlier introduction of genotype D into the Malagasy populations than the other two identified
genotypes.

It is well known that the prevalence of the different HBV genotypes and subgenotypes vary in
geographical areas and correlate strongly with ethnicity. Studying the genetic variability of the HBV
strains in endemic regions may thus be helpful when studying migration patterns of populations of
different origins. This study showed that molecular epidemiology of HBV in the North-western part of
Madagascar confirms linguistic and human genetic studies on human migrations in Africa.

**Acknowledgements**

This work was supported by a bilateral program financed by Swedish Institute scholarship (grant no:
200/01500/2007/France) and by grant for mobility Walter –Zellidja (grant no 2007-HS) and Grant
from the Swedish Research Council: VR521-2006-2573. The sequences described in this manuscript
are deposited in GenBank with accession numbers xxx- yyy.
References


Table I. Age, sex and hepatitis B serological markers of investigated inhabitants in villages with mainly farmers (Sirama) and non-farmers (Mataipako) in Northern Madagascar. Figures in parenthesis are number of individuals with anti-HBs as the only marker.

<table>
<thead>
<tr>
<th>Age</th>
<th>Males</th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>N</td>
<td>HBsAg pos</td>
<td>HBs neg</td>
<td>HBC %</td>
<td>HBC and/or anti-HBs</td>
<td>HBsAg pos</td>
<td>HBs neg</td>
<td>HBC %</td>
<td>HBC and/or anti-HBs</td>
<td>N</td>
<td>HBsAg pos</td>
<td>HBs neg</td>
<td>HBC %</td>
<td>HBC and/or anti-HBs</td>
</tr>
<tr>
<td>Sirama</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>15-19</td>
<td>27</td>
<td>3</td>
<td>19</td>
<td>81%</td>
<td></td>
<td>37</td>
<td>1</td>
<td>22 (1)</td>
<td>62%</td>
<td></td>
<td>64</td>
<td>4</td>
<td>41 (1)</td>
<td>70%</td>
</tr>
<tr>
<td>20-24</td>
<td>28</td>
<td>3</td>
<td>21</td>
<td>86%</td>
<td></td>
<td>20</td>
<td>0</td>
<td>13 (1)</td>
<td>65%</td>
<td></td>
<td>48</td>
<td>3</td>
<td>34 (1)</td>
<td>71%</td>
</tr>
<tr>
<td>25-29</td>
<td>35</td>
<td>5</td>
<td>22 (1)</td>
<td>77%</td>
<td></td>
<td>26</td>
<td>3</td>
<td>16</td>
<td>73%</td>
<td></td>
<td>61</td>
<td>8</td>
<td>38 (1)</td>
<td>62%</td>
</tr>
<tr>
<td>30-34</td>
<td>20</td>
<td>2</td>
<td>14</td>
<td>80%</td>
<td></td>
<td>32</td>
<td>3</td>
<td>16</td>
<td>59%</td>
<td></td>
<td>52</td>
<td>5</td>
<td>30</td>
<td>67%</td>
</tr>
<tr>
<td>35-39</td>
<td>20</td>
<td>2</td>
<td>11 (1)</td>
<td>65%</td>
<td></td>
<td>21</td>
<td>1</td>
<td>11</td>
<td>57%</td>
<td></td>
<td>41</td>
<td>3</td>
<td>22 (1)</td>
<td>54%</td>
</tr>
<tr>
<td>40-44</td>
<td>17</td>
<td>1</td>
<td>13</td>
<td>82%</td>
<td></td>
<td>15</td>
<td>1</td>
<td>12</td>
<td>87%</td>
<td></td>
<td>32</td>
<td>2</td>
<td>25</td>
<td>78%</td>
</tr>
<tr>
<td>45-</td>
<td>18</td>
<td>2</td>
<td>15 (1)</td>
<td>94%</td>
<td></td>
<td>6</td>
<td>1</td>
<td>5</td>
<td>100</td>
<td></td>
<td>24</td>
<td>3</td>
<td>20 (1)</td>
<td>96%</td>
</tr>
<tr>
<td>Subtotal</td>
<td></td>
<td>165</td>
<td>18</td>
<td>115 (3)</td>
<td>81%</td>
<td>157</td>
<td>10</td>
<td>95 (2)</td>
<td>67%</td>
<td></td>
<td>322</td>
<td>28</td>
<td>210 (5)</td>
<td>74%</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mataipa</td>
<td>15-19</td>
<td>21</td>
<td>1</td>
<td>4</td>
<td>24%</td>
<td>18</td>
<td>1</td>
<td>8 (1)</td>
<td>50%</td>
<td></td>
<td>39</td>
<td>2</td>
<td>12 (1)</td>
<td>36%</td>
</tr>
<tr>
<td>20-24</td>
<td>35</td>
<td>6</td>
<td>16 (1)</td>
<td>63%</td>
<td></td>
<td>26</td>
<td>2</td>
<td>15</td>
<td>65%</td>
<td></td>
<td>61</td>
<td>8</td>
<td>31 (1)</td>
<td>64%</td>
</tr>
<tr>
<td>25-29</td>
<td>19</td>
<td>1</td>
<td>11</td>
<td>63%</td>
<td></td>
<td>17</td>
<td>2</td>
<td>7</td>
<td>53%</td>
<td></td>
<td>36</td>
<td>3</td>
<td>18</td>
<td>58%</td>
</tr>
<tr>
<td>30-34</td>
<td>14</td>
<td>1</td>
<td>6</td>
<td>50%</td>
<td></td>
<td>15</td>
<td>1</td>
<td>10</td>
<td>73%</td>
<td></td>
<td>29</td>
<td>2</td>
<td>16</td>
<td>62%</td>
</tr>
<tr>
<td>35-39</td>
<td>22</td>
<td>1</td>
<td>12</td>
<td>59%</td>
<td></td>
<td>14</td>
<td>1</td>
<td>6</td>
<td>50%</td>
<td></td>
<td>36</td>
<td>2</td>
<td>18</td>
<td>55%</td>
</tr>
<tr>
<td>40-44</td>
<td>16</td>
<td>2</td>
<td>8 (1)</td>
<td>62%</td>
<td></td>
<td>7</td>
<td>1</td>
<td>1</td>
<td>29%</td>
<td></td>
<td>23</td>
<td>3</td>
<td>9 (1)</td>
<td>52%</td>
</tr>
<tr>
<td>45-</td>
<td>9</td>
<td>0</td>
<td>4 (1)</td>
<td>44%</td>
<td></td>
<td>8</td>
<td>0</td>
<td>4</td>
<td>50%</td>
<td></td>
<td>17</td>
<td>0</td>
<td>8 (1)</td>
<td>47%</td>
</tr>
<tr>
<td>Subtotal</td>
<td></td>
<td>136</td>
<td>12</td>
<td>61 (3)</td>
<td>54%</td>
<td>105</td>
<td>8</td>
<td>51 (1)</td>
<td>56%</td>
<td></td>
<td>241</td>
<td>20</td>
<td>112 (4)</td>
<td>55%</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>TOTAL</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>15-19</td>
<td>48</td>
<td>4</td>
<td>23</td>
<td>56%</td>
<td></td>
<td>55</td>
<td>2</td>
<td>30 (2)</td>
<td>58%</td>
<td></td>
<td>103</td>
<td>6</td>
<td>53 (2)</td>
<td>57%</td>
</tr>
<tr>
<td>20-24</td>
<td>63</td>
<td>9</td>
<td>37 (1)</td>
<td>73%</td>
<td></td>
<td>46</td>
<td>2</td>
<td>28 (1)</td>
<td>65%</td>
<td></td>
<td>109</td>
<td>11</td>
<td>65 (2)</td>
<td>70%</td>
</tr>
<tr>
<td>25-29</td>
<td>54</td>
<td>6</td>
<td>33 (1)</td>
<td>72%</td>
<td></td>
<td>43</td>
<td>5</td>
<td>23</td>
<td>65%</td>
<td></td>
<td>97</td>
<td>11</td>
<td>56 (1)</td>
<td>69%</td>
</tr>
<tr>
<td>30-34</td>
<td>34</td>
<td>3</td>
<td>20</td>
<td>68%</td>
<td></td>
<td>47</td>
<td>4</td>
<td>26</td>
<td>64%</td>
<td></td>
<td>81</td>
<td>7</td>
<td>46</td>
<td>65%</td>
</tr>
<tr>
<td>35-39</td>
<td>42</td>
<td>3</td>
<td>23 (1)</td>
<td>62%</td>
<td></td>
<td>35</td>
<td>2</td>
<td>17</td>
<td>54%</td>
<td></td>
<td>77</td>
<td>5</td>
<td>40 (1)</td>
<td>58%</td>
</tr>
<tr>
<td>40-44</td>
<td>33</td>
<td>3</td>
<td>21 (1)</td>
<td>73%</td>
<td></td>
<td>22</td>
<td>2</td>
<td>13</td>
<td>68%</td>
<td></td>
<td>55</td>
<td>5</td>
<td>34 (1)</td>
<td>71%</td>
</tr>
<tr>
<td>45-</td>
<td>27</td>
<td>2</td>
<td>19 (2)</td>
<td>78%</td>
<td></td>
<td>14</td>
<td>1</td>
<td>9</td>
<td>71%</td>
<td></td>
<td>41</td>
<td>3</td>
<td>28 (2)</td>
<td>76%</td>
</tr>
<tr>
<td>TOTAL</td>
<td>301</td>
<td>30</td>
<td>176 (6)</td>
<td>68%</td>
<td></td>
<td>262</td>
<td>18</td>
<td>146 (3)</td>
<td>62%</td>
<td></td>
<td>563</td>
<td>48</td>
<td>322 (9)</td>
<td>66%</td>
</tr>
</tbody>
</table>
Table II. HBeAg and anti-HBe in sera from 48 HBsAg positive inhabitants in the Sirama region and Mataipako

<table>
<thead>
<tr>
<th></th>
<th>HBsAg positive</th>
<th>HBeAg positive</th>
<th>HBeAg negative</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Male Female</td>
<td>Male Female</td>
<td>Male Female</td>
</tr>
<tr>
<td>Ambodikatakata</td>
<td>3 1</td>
<td>2 0</td>
<td>1 1</td>
</tr>
<tr>
<td>Ambodimanga*</td>
<td>6 2</td>
<td>4 0</td>
<td>2 2</td>
</tr>
<tr>
<td>Ankatoko</td>
<td>2 1</td>
<td>1 0</td>
<td>1 1</td>
</tr>
<tr>
<td>Tanambao</td>
<td>7 6</td>
<td>3 4</td>
<td>3 2</td>
</tr>
<tr>
<td>Sirama subtotal</td>
<td>18 10</td>
<td>10 4</td>
<td>7 6</td>
</tr>
<tr>
<td>Mataipako</td>
<td>12 8</td>
<td>5 1</td>
<td>7 6</td>
</tr>
<tr>
<td>Total</td>
<td>30 18</td>
<td>15 5</td>
<td>14 12</td>
</tr>
<tr>
<td>Total both sexes</td>
<td>48 20</td>
<td>26 2</td>
<td></td>
</tr>
</tbody>
</table>
**Table III. The infecting HBV genotypes in strains from 45 HBsAg positive inhabitants in five villages in Northern Madagascar**

<table>
<thead>
<tr>
<th>Genotype</th>
<th>Male/Female</th>
<th>HBsAg +</th>
<th>A/ayw1</th>
<th>A1</th>
<th>D/ayw2</th>
<th>D7</th>
<th>D/ayw3</th>
<th>D2</th>
<th>D/ayw4</th>
<th>E/ayw4</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ambodikatakata</td>
<td>3/1</td>
<td>2/0</td>
<td>1/1</td>
<td>0/0</td>
<td>0/0</td>
<td>0/0</td>
<td>0/0</td>
<td>0/0</td>
<td>0/0</td>
<td>3/1</td>
<td></td>
</tr>
<tr>
<td>Ambodimanga</td>
<td>6/2</td>
<td>3/2</td>
<td>0/0</td>
<td>0/0</td>
<td>1/0</td>
<td>1/0</td>
<td>5/2</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Tanambao</td>
<td>13</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>10</td>
<td>13</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ankatoko</td>
<td>2/1</td>
<td>0/0</td>
<td>1/0</td>
<td>0/0</td>
<td>0/0</td>
<td>0/1</td>
<td>1/1</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sirama region</td>
<td>18/10</td>
<td>5/2</td>
<td>2/3</td>
<td>1/0</td>
<td>0/0</td>
<td>1/0</td>
<td>7/5</td>
<td>16/10</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mataipako</td>
<td>12/8</td>
<td>3/0</td>
<td>1/2</td>
<td>0/0</td>
<td>0/1</td>
<td>0/0</td>
<td>8/4</td>
<td>12/7</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total M/F</td>
<td>30/18</td>
<td>8/2</td>
<td>3/5</td>
<td>1/0</td>
<td>0/1</td>
<td>1/0</td>
<td>15/9</td>
<td>28/17</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>48*</td>
<td>10</td>
<td>8</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>24</td>
<td>45</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

* Two HBV strains could not be genotyped due to low amount of HBV DNA. One individual was double infected with two strains of genotypes A and D.
FIGURE LEGEND

Figure 1
Dendrograms based on phylogenetic analysis of 681 nucleotides of the S gene in
a) UPGMA tree with 10 genotype A strains from Madagascar; b) UPGMA tree with 11 genotype D strain from Madagascar; c) Neighbor-Joining tree with 11 genotype D strains from Madagascar and d) UPGMA tree with 24 genotype E strains from Madagascar together with sequences of the corresponding region obtained from GenBank. Sequences from GenBank are indicated with accession numbers. The designation and origin of each strain are indicated at the nodes of the branches. The strains from Madagascar are indicated by arrows.
Genotype E
Genotype E