



HAL
open science

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 Magnus, Helene Norder

► **To cite this version:**

Tatiana Dupinay, Karin Restorp, Peter Leutscher, Dominique Rousset, Isabelle Chemin, et al.. High prevalence of hepatitis B virus genotype E in Northern Madagascar indicates a West-African lineage. *Journal of Medical Virology*, 2010, 82 (9), pp.1515. 10.1002/jmv.21865 . hal-00599781

HAL Id: hal-00599781

<https://hal.science/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

Journal:	<i>Journal of Medical Virology</i>
Manuscript ID:	JMV-10-1736.R1
Wiley - Manuscript type:	Research Article
Date Submitted by the Author:	28-Apr-2010
Complete List of Authors:	Dupinay, Tatiana; ISERM, U871 Restorp, Karin; Swedish Institute for Infectious Disease Control Leutscher, Peter; Aarhus University Hospital, Department of Infectious Diseases; 5 Institute Pasteur de Madagascar Rousset, Dominique; 5 Institute Pasteur de Madagascar Chemin, Isabelle; ISERM, U871 Migliani, Rene; 5 Institute Pasteur de Madagascar Magnius, Lars; Swedish Institute for Infectious Disease Control, Virological Norder, Helene; Swedish Institute for Infectious Disease Control
Keywords:	molecular epidemiology, human migration, genotyping, HBV, Africa



1
2
3
4
5 1
6
7 2
8
9
10 3
11
12 4
13
14 5
15
16 6
17
18
19 7
20
21 8
22 8 High prevalence of hepatitis B virus genotype E in Northern Madagascar
23 9 indicates a West-African lineage
24
25 10
26 11
27 12

28 13 Tatiana Dupinay^{1,2,3}, Karin Restorp⁴, Peter Leutscher^{5,6}, Dominique Rousset⁷, Isabelle
29
30 14 Chemin¹, Rene Migliani⁵, Lars Magnus⁴, Heléne Norder^{4*}.
31
32

33 15
34 16
35 17 ¹ INSERM U871, 151 Cours Albert Thomas, 69003 Lyon, France;

36
37 18 ² Université Lyon 1, IFR62 Lyon Est, 69008 Lyon, France ;

38
39
40 19 ³ Ecole pratique des hautes études , 75007 Paris France
41

42 20 ⁴ Virological Department, Swedish Institute of Infectious Disease Control, Sweden
43 21

44 22 ⁵ Institute Pasteur de Madagascar, Antananarivo, Madagascar
45 23

46 24 ⁶. Department of Infectious Diseases, Aarhus University Hospital, Skejby, DK-8200 Aarhus N,
47 25 Denmark
48 26

49
50 27 ⁷ Institute de Medecine Tropicale du Service de Sante des Armées, Marseille, France.
51 28

52
53 29 * Correspondence to: Helene Norder

54 30 Department of Virology

55 31 Swedish Institute for Infectious Disease Control, SE-171 82 Solna, Sweden

56 32 e-mail : helene.norder@smi.se
57 33
58
59 33
60

34 Word Count

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

35 Abstract : 249 words
36 Main text : 3820 words
37 4 Tables and 1 Figure
38 Running title: HBV genotype E in Madagascar

For Peer Review

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.

57 Introduction

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

69
70 Madagascar is a country with an ethnically diverse population divided into 22 tribes and a growth rate
71 of 2,8 % per year (Hewitt et al., 1996; Mafrezi and Randretsa, 1985). Madagascar is classified as an
72 area of high endemicity for HBV, and about 5% of the urban population and 30 % of the rural
73 population are chronic HBsAg carriers (Boisier *et al.*, 1996; Migliani *et al.*, 2000a; Migliani *et al.*,
74 2000b; Morvan *et al.*, 1994). Hepatitis B vaccination is the most effective strategy for prevention of
75 hepatitis B. However, in Madagascar laboratory facilities are insufficient and the use of hepatitis B
76 vaccine is limited (Migliani *et al.*, 2000b).

77
78 HBV is a small, double-stranded DNA virus and the prototype of the hepadnavirus family (Ganem &
79 Varmus, 1987). HBV strains are classified into eight genotypes, designated A-H (Arauz-Ruiz *et al.*,
80 2002; Naumann *et al.*, 1993; Norder *et al.*, 1993; Okamoto *et al.*, 1987; Stuyver *et al.*, 2000). Each
81 genotype has distinct geographical distribution. Genotype A is common in North-Western Europe,
82 North America, and Africa. Genotypes B and C are predominant in East- and South East Asia, Oceania
83 and among aboriginal populations in Australia. Genotype D is found worldwide and is predominant in
84 the Mediterranean, Middle East, and India, whereas genotype E has so far been confined to
85 populations in western parts of sub-Saharan Africa apart from one report from India (Norder et al.,
86 2004; Singh *et al.*, 2009). Genotype G has been identified in samples from several countries, and is
87 found in patients co-infected with another HBV genotype. Genotypes F and H are found mainly in
88 Central and South America, and are the indigenous genotypes of Amerindian populations (Arauz-Ruiz
89 *et al.*, 2002; Norder *et al.*, 2004; Kumagai *et al.*, 2007; Sanchez *et al.*, 2007).

1
2
3 90 The major HBV genotypes apart from E and H have been divided into subgenotypes based
4
5 91 phylogenetic analyses and genetic divergence of complete genomes. Each subgenotype has specific
6
7 92 geographical distribution. Six subgenotypes have been recognized for genotype A. A1 is prevalent in
8
9 93 Africa, Asia and Indonesia (Bowyer *et al.*, 1997; Kramvis *et al.*, 2002; Banerjee *et al.*, 2006; Kramvis
10
11 94 and Kew 2007; Tanaka *et al.*, 2004; Hannoun *et al.*, 2005). A2 is prevalent in North West Europe
12
13 95 (McMahon 2009; Norder *et al.*, 2004; Sugauchi *et al.*, 2004; Schaefer, 2007). A3 – A6 are found in
14
15 96 West and Central Africa, (Makuwa *et al.*, 2006; Kramvis and Kew 2007; Pourkarim *et al.*, 2009;
16
17 97 Kurbanov *et al.*, 2005; Olinger *et al.*, 2006). There are eight subgenotypes for genotype B, and seven
18
19 98 for C. B1 and C2 are prevalent in Japan, Korea and Northern China, C2 is also found in Alaska, while
20
21 99 B2 and C1 are prevalent in Southern China, Taiwan, and South East Asia (Norder *et al.*, 2004;
22
23 100 Schaefer, 2007; McMahon, 2009; Wang *et al.*, 2009). B3 – B5 and B8 as well as C3 – C7 are prevalent
24
25 101 in the Pacific, while B6 is found in Alaska, Northern Canada and Greenland (McMahon, 2009;
26
27 102 Mulyanto *et al.*, 2009; Cavinta *et al.*, 2009; Lusida *et al.*, 2008; Nagasaki *et al.*, 2006; Sakamoto *et al.*,
28
29 103 2006). Subgenotype B7 described from Indonesia (Nurainy *et al.*, 2008) probably represent strains of
30
31 104 subgenotype B3. There are seven subgenotypes reported for genotype D. D1 is highly prevalent in the
32
33 105 Middle East and in the Mediterranean countries, but also occurs in India and East Africa and (Banerjee
34
35 106 *et al.*, 2006; Norder *et al.*, 2004; Schaefer 2007). D2 is prevalent in East Europe including Russia and
36
37 107 the Baltic States, while D3 has a more South-Eastern distribution in Russia and is also prevalent in
38
39 108 Northern India and Pakistan (Norder *et al.*, 2004; Tallo *et al.*, 2008). D3 has also a global spread
40
41 109 attributed to its spread among injecting drug users. D5 is reported from Eastern India (Banerjee *et al.*,
42
43 110 2006; Chandra *et al.*, 2009). D4 was defined based on complete genomes from Australia and Western
44
45 111 Pacific and formed a clade with strains from North-Africa including Somalia in genetic trees based on
46
47 112 the small S gene (Norder *et al.* 2004). The North-African subclade, however, is divergent sufficiently
48
49 113 to represent a new subgenotype, D7, based on analysis of strains from Tunisia (Meldal *et al.*, 2009).
50
51 114 D6 reported from Papua (Lusida *et al.*, 2008) probably represents strains of subgenotype D4. The
52
53 115 subgenotypes of F are found in the Americas. F1 has a Northern distribution in Alaska, Mexico, and
54
55 116 Central America and has spread to Peru and Argentina, while F2 – F4 are prevalent in South America
56
57 117 (McMahon, 2009; Norder *et al.*, 2004; Devesa *et al.*, 2008).

58
59 118
60 119 In this study, the prevalence of HBV markers was investigated in populations in a rural area in the
120
121 Northern part of Madagascar. The HBV strains from chronic carriers were characterized in order to
122
123 determine the HBV genotypes and subgenotypes prevalent in this region of Madagascar.

124 **Material and Methods**

125 Materials

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

138 Serological assays

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

147 Extraction of HBV DNA, PCR amplification and sequencing of the S gene, the pre-core and the core 148 promoter regions.

149 HBV DNA was extracted from 50 µl serum by the standard phenol-chloroform method. Briefly, 50 µl
150 serum were incubated for 2 hours at 42°C with 250 mg/ml of proteinase K (Sigma, St Louis, MO, US),
151 1% sodium dodecylsulphate, 2,5 mM disodium EDTA (Sigma, St Louis, MO,US), 25 mM sodium
152 acetate and 0.25 µg/ml tRNA. The DNA was subsequently extracted with phenol and with a 1:1
153 mixture of phenol:chloroform. Thereafter the DNA was precipitated in absolute ethanol with 0.15 M of
154 potassium acetate and dissolved in 40 µl of milli-Q water.

155
156 PCR amplification of the S- region was performed by using primers hep75 and 73b in the first PCR
157 and hep3/hep33 and hep4/hep34 for nesting (Norder *et al.*, 1992). The pre-core region was amplified

1
2
3 158 with primer promC and hep68 (Sendi *et al.*, 2005), and nested with primers hep67 and hep68 (Arauz-
4 159 Ruiz *et al.*, 1997). For amplification of the core promoter region, primers hep64 and hep70 were used
5 160 in the first round of PCR and hep64 and hep54 in the nested PCR (Sendi *et al.*, 2005).
6
7 161

8
9 162 Purified products were used as templates in the sequencing reaction using the dideoxy-nucleotide
10 163 chain-termination method with an ABI PRISM BigDye terminator cycle-sequencing reaction kit
11 164 (version 3; Applied Biosystems, Foster City, CA, US). The primers used in the PCR were used as
12 165 sequencing primers. All amplification products were sequenced bi-directionally.
13
14 166 An ABI PRISM 3100 genetic analyser (Applied Biosystems) was used for electrophoresis and data
15 167 collection.
16
17 168

21 169 Phylogenetic analysis

22
23 170 Sequences obtained were edited manually using the SeqMan program in the LASERGENE package
24 171 (DNASTAR, Inc, Madison, WI). The sequences were thereafter aligned with the corresponding region
25 172 in sequences retrieved from GenBank.
26
27 173

28 174 Phylogenetic analysis was carried out with the PHYLIP program package version 3.53 (Felstenstein,
29 175 1993). Evolutionary distances were estimated with the DNADIST program using the F81 model.
30 176 Phylogenetic trees were constructed using UPGMA and neighbour-joining (N-J) algorithm in the
31 177 PHYLIP package. Genotypes and subgenotypes were determined by phylogenetic analysis of the
32 178 amplified fragments of the S gene with sequence from previously genotyped and subgenotyped strains
33 179 (Norder *et al.*, 2004). The deduced amino acid sequence of the S gene region was used to determine
34 180 the subtype, which was assessed from the substitutions at codons 122, 127, and 160 (Norder *et al.*,
35 181 1992; Okamoto *et al.*, 1987).
36
37 182

38 183 **Results**

39 184 *Seroprevalence of hepatitis B markers*

40 185 There was an overall high prevalence of HBV markers (66 %). It was significantly higher in the
41 186 Sirama villages, 74 % versus Mataipako, 55 %, ($p < 0.0001$; Fisher's exact test; Table I), whereas there
42 187 was no major difference between males and females (69% versus 65%, $p = 0.34$). In all, 370 inhabitants
43 188 had HBV markers, 322 (57 %) were negative for HBsAg but had anti-HBc and/or anti-HBs as markers
44 189 of past infection, while 48 (8.5 %) individuals were positive for HBsAg (Table I). Anti-HBs as only
45 190 serological marker was found in nine (1.6 %) inhabitants (Table I). There was a significantly higher
46 191 rate of HBV markers among males in the Sirama villages compared to Mataipako village, 70% versus

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
225

192 54%, ($p < 0.0001$; Fisher's exact test). There was no correlation between age and prevalence of HBV
193 markers (Table I).

194
195 Among the 48 HBsAg positive patients, 20 (42 %) were HBeAg positive, another 26 (54 %) were anti-
196 HBe positive, while two patients were negative for both these markers (Table II). There was no
197 significant difference in HBeAg or anti-HBe prevalence between regions or sexes.

199 *Genotype distribution*

200 Three HBV genotypes A, D, and E, were identified in 45 samples by sequencing and analyzing the S-
201 gene region (Table III). Most strains belonged to genotype E, 24/45 (53 %), followed by genotype D,
202 11/45 (24 %) and A, 10/45 (22 %). One individual was infected dually with one genotype A strain
203 expressing *ayw1* and one genotype D strain expressing *ayw2*. The HBV genotype could not be
204 determined for two individuals due to low level HBV DNA.

205
206 Most of the 24 genotype E strains originated from inhabitants in two villages, 12 from Mataipako, and
207 10 from one village in the Sirama region (Table III). Seven of the genotype A strains were from
208 inhabitants in two villages in the Sirama region, and three were from inhabitants in Mataipako. The
209 genotype D strains originated from inhabitants in all villages with HBsAg positive individuals (Table
210 III).

212 *Phylogenetic analysis*

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

220
221 All but two genotype D strains could be classified into two subgenotypes, D2 (one strain) and D7 (8
222 strains) in the phylogenetic analyses. The D7 strains, specifying *ayw2*, were found in one cluster and
223 shared similarity with strains from Somalia and Tunisia in both the UPGMA and N-J analyses (Figs
224 1b; 1c). These D7 isolates originated from inhabitants in three villages, two in the Sirama region
225 (Ambodikatakata, and Tanambao), and Mataipako. The D2 strain specifying *ayw3* was from a male

1
2
3 226 from Ankatoko in the Sirama region, and was similar to strains from East Europe (Figs 1b; 1c). The
4
5 227 D2 and D7 strains formed separate clades also in the N-J analysis, although in this analysis the D1 and
6
7 228 D3 strains were split into two and three clades, respectively (fig. 1c). The D4 and D7 strains formed
8
9 229 the first split in the N-j tree, although this was not the case in the UPGMA tree. Six D1 and two D2
10
11 230 strains classified based on complete genomes, could not be classified into a subgenotype by UPGMA
12
13 231 but belonged to their respective subgenotypes in the N-J analysis (figs 1b; 1c). On the other hand there
14
15 232 were five D1 strains based on complete genomes that were classified as D1 in the UPGMA tree but
16
17 233 could not be classified into a subgenotype by N-J (figs 1b; 1c). Two genotype D strains from
18
19 234 Madagascar were not classifiable with any of the analyses. One, 1500-02, from a female from
20
21 235 Mataipako encoding *ayw4*, was found between genotype D and E strains in the N-J tree (fig 1c). The
22
23 236 other strain, 600-02, from a male from Ambodimanga specified *ayw2*. This strain was similar to a
24
25 237 strain from aboriginal populations in Australia in the UPGMA tree and formed a separate branch
26
27 238 between the Asian D4 and the African D7 strains in the N-J tree (figs 1b; 1c).

28 239
29 240 The 24 genotype E isolates were found intermixed with strains from West Africa in the phylogenetic
30
31 241 tree (Fig. 1d). Fourteen of the Madagascar strains were found in two clades. In the major clade, 11 of
32
33 242 the Madagascar strains were intermixed with strains from Cameroon, Nigeria, Ivory Coast, and Mali
34
35 243 (Fig. 1d). Eight of these strains were from Mataipako, and four from two villages in the Sirama region,
36
37 244 Tanambao and Ankatoko. In the second cluster 4 isolates from Tanambao were intermixed with strains
38
39 245 from Congo (Fig 1d).

40 246 41 247 *Pre-core and core promoter region*

42 248 The pre-core region was sequenced in 36 strains. A premature stop codon related to the G1896A
43
44 249 mutation was observed in six sequences (16.6%). This pre-core mutant was found in three of seven
45
46 250 investigated genotype D7 strains and in three of 18 genotype E strains. It was not found in any of nine
47
48 251 genotype A strains investigated. All strains with the pre-core stop mutant were from patients with anti-
49
50 252 HBe. Another pre-core mutation, G1896T, converting Trp¹⁸⁹⁶ to Leu¹⁸⁹⁶, was found in two genotype D
51
52 253 strains and in one genotype E strain. This variant was found in two patients with anti-HBe and in one
53
54 254 patient lacking HBe markers.

55 255
56
57 256 The core promoter region was amplified and sequenced in 16 strains, three were genotype A, another
58
59 257 three were genotype D, and 10 were genotype E strains. All genotype D and E strains expressed A¹⁷⁵⁷,
60
258 and the genotype A strains expressed G¹⁷⁵⁷. One genotype A strain had C1766T and T1768A

1
2
3 259 mutations in the core promoter region, and one genotype E strain had a triple mutation, A1762C,
4 260 G1764T and C1766G. All other strains had wild type promoter sequence.
5
6

7 261 8 9 262 **Discussion**

10 263 The prevalence of hepatitis B markers in the rural populations from the Northern region of Madagascar
11 264 was high, although not as high as the prevalence described in other rural population in the centre of
12 265 Madagascar; where up to 18 % were HBsAg positive (Boisier *et al.*, 1996; Migliani *et al.*, 2000a;
13 266 Morvan *et al.*, 1994). Only 1.6 % of the inhabitants from this study had anti-HBs only which strongly
14 267 suggest that implementation of control programs for hepatitis B in these areas is warranted. The
15 268 present findings on the epidemiological features of persons infected with hepatitis B such as age, sex,
16 269 and origin showed that the prevalence of infection appears to be higher among males from rural
17 270 villages.
18
19
20
21
22
23
24

25 271
26 272 In this study, there was a high frequency of HBeAg positive HBsAg carriers (42 %), mainly among
27 273 those infected with genotypes A1 and E. This prevalence of HBeAg was higher than that from other
28 274 parts of East and Central Africa as Zimbabwe, Cameroon, Ethiopia, Benin and the Central African
29 275 Republic (Tswana *et al.*, 1996; Rapicetta *et al.*, 1991; Abebe *et al.*, 2003; Fujiwara *et al.*, 2005;
30 276 Pawlowsky *et al.*, 1995; Abiodun *et al.*, 1994). In addition, there was a low prevalence of strains with
31 277 mutations in the pre-core and core promoter despite the relatively high presence of A1 and D strains.
32 278 In Mediterranean countries, genotype D has been shown to present in high frequency in negative
33 279 chronic HBV infection without HBeAg, associated with HBV mutants in the pre-core region (Brunetto
34 280 *et al.*, 1993). The most common precore mutation, G1896A, which creates a stop codon, has been
35 281 associated previously with the subgenotypes found in this study (Chu *et al.*, 2002; Kramvis *et al.*,
36 282 1997; Lok *et al.*, 1994; Olinger *et al.*, 2006). This mutant was only found in 38 % of the genotype D
37 283 strains and in none of the A1 strains. There were also only two strains, one genotype A and one
38 284 genotype E, with core promoter mutations. The high HBeAg prevalence and low frequency of precore
39 285 and core promoter may be due to the prevalent HBV genotypes if genotypes A1 and E have prolonged
40 286 replicative phase with HBeAg expression, as is described for genotypes B and C in Asia (Kao, 2002;
41 287 Orito *et al.*, 2001). It is not known if chronic carriers infected with A1 and E have a higher prevalence
42 288 of HBeAg positivity than those infected with genotype D also in other parts of Africa, or if it is
43 289 specific for this studied population.
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59

60 291 The Malagasy population has a widespread origin from Asiatic to sub-Saharan African descendants
292 and a mixture of the two (Hurles *et al.*, 2005; Singer *et al.*, 1957). Although both African and

1
2
3 293 Southeast Asian populations have contributed to Madagascar's gene pool, those of Asian origin are
4 294 most predominant in the central highlands, while the coastal populations often are of African origin
5 295 (Chow *et al.*, 2005; Regueiro *et al.*, 2008; Hurles *et al.*, 2005).
6
7 296
8
9 297 The origin of the population may also be reflected by the prevalent HBV genotypes. The subgenotype
10 298 A1 strains were similar and mainly found in two villages. This subgenotype dominates in all East
11 299 African countries including South Africa, Malawi, Tanzania, Uganda, and Somalia (Kramvis & Kew,
12 300 2007). The isolates from Madagascar were similar to Malawi isolates, suggesting that this subgenotype
13 301 was introduced into Madagascar from East Africa. Geographically Malawi is close to the North
14 302 Western part of Madagascar and historically, in the nineteenth century, the trade route from Malawi to
15 303 the western coast of Madagascar became exploited increasingly for slave trading.
16 304
17
18 305 The genotype E strains in this study showed high sequence similarity to West African strains, where
19 306 this genotype is confined. Previous studies have shown a low genetic diversity of genotype E strains
20 307 isolated from a large geographical area in West Africa (Kramvis *et al.*, 2005; Quintero *et al.*, 2002).
21 308 Several studies have proposed that this low genetic divergence together with few findings of genotype
22 309 E in the Americas indicates a recent introduction of genotype E into the human population (Mulders *et*
23 310 *al.*, 2004; Odemuyiwa *et al.*, 2001; Suzuki *et al.*, 2003). The genotype E strains in this study were
24 311 mainly from inhabitants in two villages, which may indicate a recent introduction, and that sufficient
25 312 time has not yet passed for its subsequent spread into the rest of the country. The inhabitants in the
26 313 two villages were working in the rice fields, and they may be slave descendant with origin from
27 314 mainland Africa. West African origin of some Malagasy populations is supported by genetic studies on
28 315 Y-chromosomal polymorphism and mitochondrial sequence diversities and on β -globulin haplotypes
29 316 (Hurles *et al.*, 2005; Hewitt *et al.*, 1996). In addition, hepatitis C virus genotype 2 strains, common in
30 317 West Africa, have been isolated from Malagasy populations, further supporting previous human
31 318 migrations from West Africa to Madagascar (Markov *et al.*, 2009).
32 319
33
34 320 Interestingly, 24 % of the strains in this study were genotype D, and found in all five villages with
35 321 HBsAg positive individuals. The majority were subgenotype D7, which are also found in Somalia and
36 322 Tunisia. This is consistent with the subdivision of the previous classified D4 strains into two
37 323 subgenotypes, D4 represented by strains from the Australia and Western Pacific, and D7 with North
38 324 African strains (Meldal *et al.* 2009; Schaefer *et al.*, 2009). The D4 and D7 strains formed the first split
39 325 in the genotype D phylogenetic tree, when N-J algorithm was used, as they do when complete
40 326 genomes are analysed (Meldal *et al.*, 2009; Norder *et al.*, 2004; Tallo *et al.*, 2008). N-J analyses of

1
2
3 327 genotype D S –genes thus seems better to mirror the tree obtained from complete genomes than
4
5 328 UPGMA. The split of genotype D strains into two main branches, one with D4 and D7 and one with
6
7 329 the other D subgenotypes, with D5 as first spit, might support that genotype D has followed two waves
8
9 330 of two human migrations out of Africa. The D4 subgenotype being linked to the early settlement of
10
11 331 Papua-Australia-Melanesia followed by a later wave of settlement reaching India supported by India,
12 332 being populated before Europe, and D5 being the first split in this branch.
13

14 333
15
16 334 There were in this study five different D7 strains, all showing close relationship with strains from
17
18 335 Somalia. It has been proposed that there has been a migration wave into Madagascar from central East
19 336 Africa based on human genetic and linguistic studies (Dahl, 1988; Hewitt *et al.*, 1996; Hurles *et al.*,
20
21 337 2005). Interestingly, there was one genotype D strain similar showing a distant genetic relationship to
22
23 338 a strain from Australia, and formed a separate branch between the African D7 and the Asian D4
24
25 339 strains, suggesting also Asian import of genotype D into Madagascar. This is in agreement with
26
27 340 genetic studies showing a South East Asian and Pacific origin of some Malagasy populations (Hurles
28 341 *et al.*, 2005; Migot *et al.*, 1995; Soodyall *et al.*, 1995). The dissemination of genotype D in the
29
30 342 majority of the villages and its higher genetic variability compared to genotype A and E, indicate an
31
32 343 earlier introduction of genotype D into the Malagasy populations than the other two identified
33
34 344 genotypes.
35

36 345
37 346 It is well known that the prevalence of the different HBV genotypes and subgenotypes vary in
38
39 347 geographical areas and correlate strongly with ethnicity. Studying the genetic variability of the HBV
40
41 348 strains in endemic regions may thus be helpful when studying migration patterns of populations of
42
43 349 different origins. This study showed that molecular epidemiology of HBV in the North-western part of
44 350 Madagascar confirms linguistic and human genetic studies on human migrations in Africa.
45

46 351 47 48 352 **Acknowledgements**

49 353 This work was supported by a bilateral program financed by Swedish Institute scholarship (grant no:
50 354 200/01500/2007/France) and by grant for mobility Walter –Zellidja (grant no 2007-HS) and Grant
51 355 from the Swedish Research Council: VR521-2006-2573. The sequences described in this manuscript
52
53 356 are deposited in GenBank with accession numbers xxx- yyy.
54
55
56
57
58
59
60

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
404
405357 **References**

- 358 Abebe A, Nokes DJ, Dejene A, Enquselassie F, Messele T, Cutts F T. 2003. Seroepidemiology of
359 hepatitis B virus in Addis Ababa, Ethiopia: transmission patterns and vaccine control.
360 *Epidemiol Infect* 131:757-770.
- 361 Abiodun PO, Olomu A, Okolo SN, Obasohan, A, Freeman O. 1994. The prevalence of hepatitis Be
362 antigen and anti-HBe in adults in Benin City. *West Afr J Med* 13:171-174.
- 363 Arauz-Ruiz P, Norder H, Robertson BH, Magnius LO. 2002. Genotype H: a new Amerindian genotype
364 of hepatitis B virus revealed in Central America. *J Gen Virol* 83:2059-2073.
- 365 Arauz-Ruiz P, Norder H, Visona KA, Magnius LO. 1997. Genotype F prevails in HBV infected
366 patients of hispanic origin in Central America and may carry the precore stop mutant. *J Med
367 Virol* 51:305-312.
- 368 Banerjee A, Kurbanov F, Datta S, Chandra PK, Tanaka Y, Mizokami M, Chakravarty R. 2006.
369 Phylogenetic relatedness and genetic diversity of hepatitis B virus isolates in Eastern India. *J
370 Med Virol* 78:1164-1174.
- 371 Boisier P, Rabarijaona L, Piollet M, Roux JF, Zeller HG. 1996. Hepatitis B virus infection in general
372 population in Madagascar: evidence for different epidemiological patterns in urban and in rural
373 areas. *Epidemiol Infect* 117:133-137.
- 374 Bowyer SM, van Staden L, Kew MC, Sim JG. 1997. A unique segment of the hepatitis B virus group
375 A genotype identified in isolates from South Africa. *J Gen Virol* 78:1719-1729.
- 376 Brunetto MR, Giarin M, Saracco G, Oliveri F, Calvo P, Capra G, Randone A, Abate ML, Manzini P,
377 Capalbo M. 1993. Hepatitis B virus unable to secrete e antigen and response to interferon in
378 chronic hepatitis B. *Gastroenterology* 105:845-850.
- 379 Cavinta L, Sun J, May A, Yin J, von Meltzer M, Radtke M, Barzaga NG, Cao G, Schaefer S. 2009. A
380 new isolate of hepatitis B virus from the Philippines possibly representing a new subgenotype
381 C6. *J Med Virol*. 81, 983-7
- 382 Chandra PK, Biswas A, Datta S, Banerjee A, Panigrahi R, Chakrabarti S, De BK, Chakravarty R.
383 2009. Subgenotypes of hepatitis B virus genotype D (D1, D2, D3 and D5) in India: differential
384 pattern of mutations, liver injury and occult HBV infection. *J Viral Hepat*. 16: 749-56.
- 385 Chow RA, Caeiro JL, Chen SJ, Garcia-Bertrand RL, Herrera RJ. 2005. Genetic characterization of
386 four Austronesian-speaking populations. *J Hum Genet*. 50:550-9.
- 387 Chu CM, Yeh CT, Lee CS, Sheen IS, Liaw YF. 2002. Precore stop mutant in HBeAg-positive patients
388 with chronic hepatitis B: clinical characteristics and correlation with the course of HBeAg-to-
389 anti-HBe seroconversion. *J Clin Microbiol* 40:16-21.
- 390 Dahl OC. 1988. Bantu substratum in Malagasy. *Etudes Océan Indien* 9:91-132.
- 391 Devesa M, Loureiro CL, Rivas Y, Monsalve F, Cardona N, Duarte MC, Pobleto F, Gutierrez MF,
392 Botto C, Pujol FH. 2008. Subgenotype diversity of hepatitis B virus American genotype F in
393 Amerindians from Venezuela and the general population of Colombia. *J Med Virol*. 80: 20-6.
- 394 Felstenstein J. 1993. PHYLIP : phylogeni inference package, Version3.52c. Seattle, WA : University
395 of Washington.
- 396 Fujiwara K, Tanaka Y, Orito E, Ohno T, Kato T, Sugihara K, Hasegawa I, Sakurai M, Ito K, Ozasa A,
397 Sakamoto Y, Arita I, El-Gohary A, Benoit A, Ogroundele-Akplogan SI, Yoshihara N, Ueda R,
398 Mizokami M. 2005. Distribution of HBV genotypes among HBV carriers in
399 Benin: phylogenetic analysis and virological characteristics of HBV genotype E. *World J
400 Gastroenterol* 11: 6410-6415.
- 401 Ganem D, Varmus HE. 1987. The molecular biology of the hepatitis B viruses. *Annu Rev Biochem*
402 56:651-693.
- 403 Hannoun C, Soderstrom A, Norkrans G, Lindh M. 2005. Phylogeny of African complete genomes
404 reveals a West African genotype A subtype of hepatitis B virus and relatedness between Somali
405 and Asian A1 sequences. *J Gen Virol* 86:2163-2167.

1

2

3

4

5

6

7

8

9

10

11

12

13

14

15

16

17

18

19

20

21

22

23

24

25

26

27

28

29

30

31

32

33

34

35

36

37

38

39

40

41

42

43

44

45

46

47

48

49

50

51

52

53

54

55

56

57

58

59

60

- 406 Heikel J, Sekkat S, Bouqdir F, Rich H, Takourt B, Radouani F, Hda N, Ibrahimy S, Benslimane A.
407 1999. The prevalence of sexually transmitted pathogens in patients presenting to a Casablanca
408 STD clinic. *Eur J Epidemiol* 15 :711-715.
- 409 Hewitt R, Krause A, Goldman A, Campbell G, Jenkins T. 1996. Beta-globin haplotype analysis
410 suggests that a major source of Malagasy ancestry is derived from Bantu-speaking Negroids.
411 *Am J Hum Genet.* 58:1303-8
- 412 Hurles ME, Sykes BC, Jobling MA, Forster P. 2005. The dual origin of the Malagasy in Island
413 Southeast Asia and East Africa: evidence from maternal and paternal lineages. *Am J Hum*
414 *Genet* 76:894-901.
- 415 Kao JH. 2002. Hepatitis B viral genotype: Clinical relevance and molecular characteristics. *J.*
416 *Gastroenterol Hepatol* 17:643-650.
- 417 Kramvis A, Bukofzer S, Kew MC, Song E. 1997. Nucleic acid sequence analysis of the precore region
418 of hepatitis B virus from sera of southern African black adult carriers of the virus. *Hepatology*
419 25:235-240.
- 420 Kramvis A, Kew MC. 2007. Molecular characterization of subgenotype A1 (subgroup Aa) of hepatitis
421 B virus. *Hepatol Res* 37:S27-32.
- 422 Kramvis A, Restorp K, Norder H, Botha JF, Magnius LO, Kew MC. 2005. Full genome analysis of
423 hepatitis B virus genotype E strains from South-Western Africa and Madagascar reveals low
424 genetic variability. *J Med Virol* 77:47-52.
- 425 Kramvis A, Weitzmann L, Owiredu WK, Kew MC. 2002. Analysis of the complete genome of
426 subgroup A' hepatitis B virus isolates from South Africa. *J Gen Virol* 83:835-839.
- 427 Kumagai I, Abe K, Oikawa T, Sato A, Sato S, Endo R, Takikawa Y, Suzuki K, Masuda T, Sainokami
428 S, Endo K, Takahashi M, Okamoto H. 2007. A male patient with severe acute hepatitis who
429 was domestically infected with a genotype H hepatitis B virus in Iwate, Japan. *J Gastroenterol*
430 42:168-175.
- 431 Kurbanov F, Tanaka Y, Fujiwara K, Sugauchi F, Mbanya D, Zekeng L, Ndembi N, Ngansop C,
432 Kaptue L, Miura T, Ido E, Hayami M, Ichimura H, Mizokami M. 2005. A new subtype
433 (subgenotype) Ac (A3) of hepatitis B virus and recombination between genotypes A and E in
434 Cameroon. *J Gen Virol* 86:2047-2056.
- 435 Leutscher PD, Ramarokoto CE, Hoffmann S, Jensen JS, Ramaniraka V, Randrianasolo B, Raharisolo
436 C, Migliani R, Christensen N. 2008. Coexistence of urogenital schistosomiasis and sexually
437 transmitted infection in women and men living in an area where *Schistosoma haematobium* is
438 endemic. *Clin Infect Dis* 47 :775-782.
- 439 Lok AS, Akarca U, Greene S. 1994. Mutations in the pre-core region of hepatitis B virus serve to
440 enhance the stability of the secondary structure of the pre-genome encapsidation signal. *Proc*
441 *Natl Acad Sci U S A* 91:4077-4081.
- 442 Lusida MI, Nugrahaputra VE, Soetjipto Handajani R, Nagano-Fujii M, Sasayama M, Utsumi T, Hotta
443 H. 2008. Novel subgenotypes of hepatitis B virus genotypes C and D in Papua, Indonesia. *J*
444 *Clin Microbiol.* 46:2160-6.
- 445 Mafrezi A, Randretsa I. 1985. Evolution and structure of the population of Madagascar. *Demogr Afr.*
446 48-49:51-60.
- 447 Makuwa, M., Souquiere, S., Telfer, P., Apetrei, C., Vray, M., Bedjabaga, I., Mouinga-Ondeme, A.,
448 Onanga, R., Marx, P. A., Kazanji, M., Roques, P. & Simon, F. (2006). Identification of
449 hepatitis B virus subgenotype A3 in rural Gabon. *J Med Virol* 78, 1175-1184.
- 450 Markov PV, Pepin J, Frost E, Deslandes S, Labbe AC, Pybus OG. 2009. Phylogeography and
451 molecular epidemiology of hepatitis C virus genotype 2 in Africa. *J Gen Virol* 90:2086-2096.
- 452 McMahan BJ. 2009. The influence of hepatitis B virus genotype and subgenotype on the natural
453 history of chronic hepatitis B. *Hepatol Int.* 3:334-42.
- 454 Meldal BH, Moula NM, Barnes IH, Boukef K, Allain JP. 2009. A novel hepatitis B virus subgenotype,
455 D7, in Tunisian blood donors. *J Gen Virol.* 90:1622-8.

- 1
2
3 456 Migliani R, Rakoto Andrianarivelo M, Rousset D, Rabarijaona L, Randrianarisoa P, Roux J F. 2000a.
4 457 Seroprevalence of viral hepatitis B in the city of Mahajanga, Madagascar in 1999. *Med Trop*
5 458 (Mars) 60 :146-150.
- 7 459 Migliani R, Rousset D, Rakoto-Andrianarivelo M, Rabarijaona L, Ratsitorahina M, Rajaonarivony V,
8 460 Maucelere P. 2000b. Hepatitis B virus infection: a public health problem in Madagascar. *Arch*
9 461 *Inst Pasteur Madagascar* 66:50-54.
- 10 462 Migot F, Perichon B, Danze PM, Raharimalala L, Lepers JP, Deloron P, Krishnamoorthy R. 1995.
11 463 HLA class II haplotype studies bring molecular evidence for population affinity between
12 464 Madagascans and Javanese. *Tissue Antigens*. 46:131-5.
- 14 465 Morvan JM, Boisier P, Andrianimanana D, Razainirina J, Rakoto-Andrianarivelo M, Roux JF. 1994.
15 466 Serological markers for hepatitis A, B and C in Madagascar. First investigation in a rural area.
16 467 *Bull Soc Pathol Exot* 87:138-142.
- 18 468 Mulders MN, Venard V, Njayou M, Edoth AP, Bola Oyefolu AO, Kehinde MO, Muyembe Tamfum
19 469 JJ, Nebie YK, Maiga I, Ammerlaan W, Fack F, Omilabu SA, Le Faou A, Muller CP. 2004.
20 470 Low genetic diversity despite hyperendemicity of hepatitis B virus genotype E throughout
21 471 West Africa. *J Infect Dis* 190 :400-408.
- 22 472 Mulyanto Depamede SN, Surayah K, Tsuda F, Ichiyama K, Takahashi M, Okamoto, H. 2009. A
23 473 nationwide molecular epidemiological study on hepatitis B virus in Indonesia: identification of
24 474 two novel subgenotypes, B8 and C7. *Arch Virol*. 154 :1047-59.
- 26 475 Nagasaki F, Niitsuma H, Cervantes JG, Chiba M, Hong S, Ojima T, Ueno Y, Bondoc E, Kobayashi K,
27 476 Ishii M, Shimosegawa T. 2006. Analysis of the entire nucleotide sequence of hepatitis B virus
28 477 genotype B in the Philippines reveals a new subgenotype of genotype B. *J Gen Virol* 87:1175-
29 478 1180.
- 31 479 Naumann H, Schaefer S, Yoshida CF, Gaspar AM, Repp R, Gerlich WH. 1993. Identification of a
32 480 new hepatitis B virus (HBV) genotype from Brazil that expresses HBV surface antigen subtype
33 481 adw4. *J Gen Virol* 74:1627-1632.
- 34 482 Norder H, Courouce AM, Coursaget P, Echevarria JM, Lee SD, Mushahwar IK, Robertson BH,
35 483 Locarnini S, Magnius LO. 2004. Genetic diversity of hepatitis B virus strains derived
36 484 worldwide: genotypes, subgenotypes, and HBsAg subtypes. *Intervirology* 47:289-309.
- 38 485 Norder H, Courouce AM, Magnius LO. 1992. Molecular basis of hepatitis B virus serotype variations
39 486 within the four major subtypes. *J Gen Virol* 73:3141-3145.
- 40 487 Norder H, Courouce AM, Magnius LO. 1994. Complete genomes, phylogenetic relatedness, and
41 488 structural proteins of six strains of the hepatitis B virus, four of which represent two new
42 489 genotypes. *Virology* 198:489-503.
- 44 490 Nurainy N, Muljono DH, Sudoyo H, Marzuki S. 2008. Genetic study of hepatitis B virus in Indonesia
45 491 reveals a new subgenotype of genotype B in east Nusa Tenggara. *Arch Virol*. 153: 1057-65.
- 46 492 Odemuyiwa SO, Mulders MN, Oyedele OI, Ola SO, Odaibo GN, Olaleye DO, Muller CP. 2001.
47 493 Phylogenetic analysis of new hepatitis B virus isolates from Nigeria supports endemicity of
48 494 genotype E in West Africa. *J Med Virol* 65:463-469.
- 49 495 Okamoto H, Imai M, Tsuda F, Tanaka T, Miyakawa Y, Mayumi M. 1987. Point mutation in the S gene
50 496 of hepatitis B virus for a d/y or w/r subtypic change in two blood donors carrying a surface
51 497 antigen of compound subtype ady or adwr. *J Virol* 61:3030-3034.
- 53 498 Olinger CM, Venard V, Njayou M, Oyefolu AO, Maiga I, Kemp AJ, Omilabu SA, le Faou A, Muller
54 499 CP. 2006. Phylogenetic analysis of the precore/core gene of hepatitis B virus genotypes E and
55 500 A in West Africa: new subtypes, mixed infections and recombinations. *J Gen Virol* 87:1163-
56 501 1173.
- 58 502 Orito E, Mizokami M, Sakugawa H, Michitaka K, Ishikawa K, Ichida T, Okanoue T, Yotsuyanagi H,
59 503 Iino S. 2001. A case-control study for clinical and molecular biological differences between
60 504 hepatitis B viruses of genotypes B and C. *Hepatology* 33:218-223.

- 1
2
3 505 Pawlotsky JM, Belec L, Gresenguet G, Deforges L, Bouvier M, Duval J, Dhumeaux D. 1995. High
4 506 prevalence of hepatitis B, C, and E markers in young sexually active adults from the Central
5 507 African Republic. *J Med Virol* 46:269-272.
- 7 508 Pourkarim MR, Lemey P, Amini-Bavil-Olyaei S, Maes P, Van Ranst M. 2009. Novel hepatitis B virus
8 509 subgenotype A6 in African-Belgian patients. *J Clin Virol*. In press.
- 9 510 Quintero A, Martinez D, Alarcon De Noya B, Costagliola A, Urbina L, Gonzalez N, Liprandi F,
10 511 Castro De Guerra D, Pujol FH. 2002. Molecular epidemiology of hepatitis B virus in Afro-
11 512 Venezuelan populations. *Arch Virol* 147:1829-1836.
- 12 513 Ramarakoto CE, Leutscher PD, van Dam G, Christensen NO. 2008. Ultrasonographical findings in the
13 514 urogenital organs in women and men infected with *Schistosoma haematobium* in northern
14 515 Madagascar. *Trans R Soc Trop Med Hyg* 102 :767-773.
- 15 516 Rapicetta M, Stroffolini T, Ngatchu T, Chionne P, Ciccaglione AR, Lantum D, Chiaramonte M. 1991.
16 517 Age- and sex-related study of HBV-DNA in HBsAg asymptomatic children from an endemic
17 518 area (Cameroon). *Ann Trop Paediatr* 11 :325-329.
- 18 519 Regueiro M, Mirabal S, Lacau H, Caeiro JL, Garcia-Bertrand RL, Herrera RJ. 2008. Austronesian
19 520 genetic signature in East African Madagascar and Polynesia. *J Hum Genet*. 53: 106-20.
- 20 521 Sakamoto T, Tanaka Y, Orito E, Co J, Clavio J, Sugauchi F, Ito K, Ozasa A, Quino A, Ueda R,
21 522 Sollano J, Mizokami M. 2006. Novel subtypes (subgenotypes) of hepatitis B virus genotypes B
22 523 and C among chronic liver disease patients in the Philippines. *J Gen Virol* 87:1873-1882.
- 23 524 Sanchez LV, Tanaka Y, Maldonado M, Mizokami M, Panduro A. 2007. Difference of hepatitis B virus
24 525 genotype distribution in two groups of mexican patients with different risk factors. High
25 526 prevalence of genotype H and G. *Intervirology* 50:9-15.
- 26 527 Schaefer S. 2007. Hepatitis B virus taxonomy and hepatitis B virus genotypes. *World J Gastroenterol*
27 528 13:14-21.
- 28 529 Schaefer S, Magnius L, Norder H. 2009. Under Construction: Classification of Hepatitis B Virus
29 530 Genotypes and Subgenotypes. *Intervirology*. 52:323-325.
- 30 531 Sendi H, Mehrab-Mohseni M, Zali MR, Norder H, Magnius LO. 2005. T1764G1766 core promoter
31 532 double mutants are restricted to Hepatitis B virus strains with an A1757 and are common in
32 533 genotype D. *J Gen Virol* 86:2451-2458.
- 33 534 Singer R, Budtz-Olsen OE, Brain P, Saugrain J. 1957. Physical features, sickling and serology of the
34 535 Malagasy of Madagascar. *Am J Phys Anthropol* 15:91-124.
- 35 536 Singh J, Dickens C, Pahal V, Kumar R, Chaudhary R, Kramvis A, Kew MC. 2009. First report of
36 537 genotype e of hepatitis B virus in an Indian population. *Intervirology* 52:235-238.
- 37 538 Soodyall H, Jenkins T, Stoneking M. 1995. 'Polynesian' mtDNA in the Malagasy. *Nat Genet*. 10: 377-
38 539 8.
- 39 540 Stuyver L, De Gendt S, Van Geyt C, Zoulim F, Fried M, Schinazi RF, Rossau R. 2000. A new
40 541 genotype of hepatitis B virus: complete genome and phylogenetic relatedness. *J Gen Virol* 81:
41 542 67-74.
- 42 543 Sugauchi F, Kumada H, Acharya SA, Shrestha SM, Gamutan MT, Khan M, Gish RG, Tanaka Y, Kato
43 544 T, Orito E, Ueda R, Miyakawa Y, Mizokami M. 2004. Epidemiological and sequence
44 545 differences between two subtypes (Ae and Aa) of hepatitis B virus genotype A. *J Gen Virol*
45 546 85:811-820.
- 46 547 Suzuki S, Sugauchi F, Orito E, Kato H, Usuda S, Siransy L, Arita L, Sakamoto Y, Yoshihara N, El-
47 548 Gohary A, Ueda R, Mizokami M. 2003. Distribution of hepatitis B virus (HBV) genotypes
48 549 among HBV carriers in the Cote d'Ivoire: complete genome sequence and phylogenetic
49 550 relatedness of HBV genotype E. *J Med Virol* 69:459-465.
- 50 551 Tallo T, Tefanova V, Priimägi L, Schmidt J, Katargina O, Michailov M, Mukomolov S, Magnius L,
51 552 Norder H. 2008. D2: major subgenotype of hepatitis B virus in Russia and the Baltic region. *J*
52 553 *Gen Virol*. 89:1829-39.

- 1
2
3 554 Tanaka Y, Hasegawa I, Kato T, Orito E, Hirashima N, Acharya SK, Gish RG, Kramvis A, Kew MC,
4 555 Yoshihara N, Shrestha SM, Khan M, Miyakawa Y, Mizokami M. 2004. A case-control study
5 556 for differences among hepatitis B virus infections of genotypes A (subtypes Aa and Ae) and D.
6 557 Hepatology 40:747-55.
- 8 558 Tanaka Y, Orito E, Yuen MF, Mukaide M, Sugauchi F, Ito K, Ozasa A, Sakamoto T, Kurbanov F, Lai
9 559 CL, Mizokami M. 2005. Two subtypes (subgenotypes) of hepatitis B virus genotype C: A
10 560 novel subtyping assay based on restriction fragment length polymorphism. Hepatol Res 33:
11 561 216-224.
- 13 562 Toukan AU. 1996. Hepatitis B in the Middle East: aspects of epidemiology and liver disease after
14 563 infection. Gut 38:S2-4.
- 15 564 Tswana S, Chetsanga C, Nyström L, Moyo S, Nzara M, Chieza LA. 1996. Sero-epidemiological cross-
16 565 sectional study of hepatitis B virus in Zimbabwe. S Afr Med J 86:72-5. WHO Hepatitis B. Fact
17 566 sheet N°204.
- 18 567 Wang HY, Li D, Liu W, Jin X, Du B, Li YP, Gu HX, Zhang SY. 2009. Hepatitis B virus subgenotype
19 568 C2 is the most prevalent subgenotype in northeast China. Clin Microbiol Infect. In Press.
- 21 569
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

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.

Age	<i>Males</i>				<i>Females</i>				<i>TOTAL</i>			
	N	HBsAg pos	HBsAg neg anti-HBc and/or anti-HBs	% with total markers	N	HBsAg pos	HBsAg neg anti-HBc and/or anti-HBs	% with total markers	N	HBsAg pos	HBsAg neg anti-HBc and/or anti-HBs	% with total markers
Sirama												
15-19	27	3	19	81%	37	1	22 (1)	62%	64	4	41 (1)	70%
20-24	28	3	21	86%	20	0	13 (1)	65%	48	3	34 (1)	71%
25-29	35	5	22 (1)	77%	26	3	16	73%	61	8	38 (1)	62%
30-34	20	2	14	80%	32	3	16	59%	52	5	30	67%
35-39	20	2	11 (1)	65%	21	1	11	57%	41	3	22 (1)	54%
40-44	17	1	13	82%	15	1	12	87%	32	2	25	78%
45-	18	2	15 (1)	94%	6	1	5	100%	24	3	20 (1)	96%
Sub-total	165	18	115 (3)	81%	157	10	95 (2)	67%	322	28	210 (5)	74%
		10.9%	70%			6.4%	60%			8.7%	65%	
Mataipako												
15-19	21	1	4	24%	18	1	8 (1)	50%	39	2	12 (1)	36%
20-24	35	6	16 (1)	63%	26	2	15	65%	61	8	31 (1)	64%
25-29	19	1	11	63%	17	2	7	53%	36	3	18	58%
30-34	14	1	6	50%	15	1	10	73%	29	2	16	62%
35-39	22	1	12	59%	14	1	6	50%	36	2	18	55%
40-44	16	2	8 (1)	62%	7	1	1	29%	23	3	9 (1)	52%
45-	9	0	4 (1)	44%	8	0	4	50%	17	0	8 (1)	47%
Sub-total	136	12	61 (3)	54%	105	8	51 (1)	56%	241	20	112 (4)	55%
		8.8%	45%			7.6%				8.3%	46%	
TOTAL												
15-19	48	4	23	56%	55	2	30 (2)	58%	103	6	53 (2)	57%
20-24	63	9	37 (1)	73%	46	2	28 (1)	65%	109	11	65 (2)	70%
25-29	54	6	33 (1)	72%	43	5	23	65%	97	11	56 (1)	69%
30-34	34	3	20	68%	47	4	26	64%	81	7	46	65%
35-39	42	3	23 (1)	62%	35	2	17	54%	77	5	40 (1)	58%
40-44	33	3	21 (1)	73%	22	2	13	68%	55	5	34 (1)	71%
45-	27	2	19 (2)	78%	14	1	9	71%	41	3	28 (2)	76%
TOTAL	301	30	176 (6)	68%	262	18	146 (3)	62%	563	48	322 (9)	66%
		10%				6.7%				8.5%		

Table II. HBeAg and anti-HBe in sera from 48 HBsAg positive inhabitants in the Sirama region and Mataipako

	HBsAg positive		HBeAg positive Anti-HBe negative		HBeAg negative Anti-HBe positive		HBeAg negative Anti-HBe negative	
	Male	Female	Male	Female	Male	Female	Male	Female
Ambodikatakata	3	1	2	0	1	1	0	0
Ambodimanga*	6	2	4	0	2	2	0	0
Ankatoko	2	1	1	0	1	1	0	0
Tanambao	7	6	3	4	3	2	1	0
Sirama subtotal	18	10	10	4	7	6	1	0
Mataipako	12	8	5	1	7	6	0	1
Total	30	18	15	5	14	12	1	1
Total both sexes	48		20		26		2	

Table III. The infecting HBV genotypes in strains from 45 HBsAg positive inhabitants in five villages in Northern Madagascar

	HBsAg +	Genotype Male/ Female						Total
		A/ayw1 A1	D/ayw2 D7	D/ayw3 D2	D/ayw4	D/ayw2	E/ayw4	
Ambodikatakata	3/1	2/0	1/1	0/0	0/0	0/0	0/0	3/1
Subtotal	4	2	2	0	0	0	0	4
Ambodimanga	6/2	3/2	0/0	0/0	0/0	1/0	1/0	5/2
Subtotal	8	5	0	0	0	1	1	7
Tanambao	7/6	0/0	1/2	0/0	0/0	0/0	6/4	7/6
Subtotal	13	0	3	0	0	0	10	13
Ankatoko	2/1	0/0	0/0	1/0	0/0	0/0	0/1	1/1
Subtotal	3	0	0	1	0	0	1	2
Subtotal	18/10	5/2	2/3	1/0	0/0	1/0	7/5	16/10
Sirama regioion	28	7	5	1	0	1	12	26
Mataipako	12/8	3/0	1/2	0/0	0/1	0/0	8/4	12/7
Subtotal	20	3	3	0	1	0	12	19
Total M/F	30/18	8/2	3/5	1/0	0/1	1/0	15/9	28/17
Total	48*	10	8	1	1	1	24	45

* 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

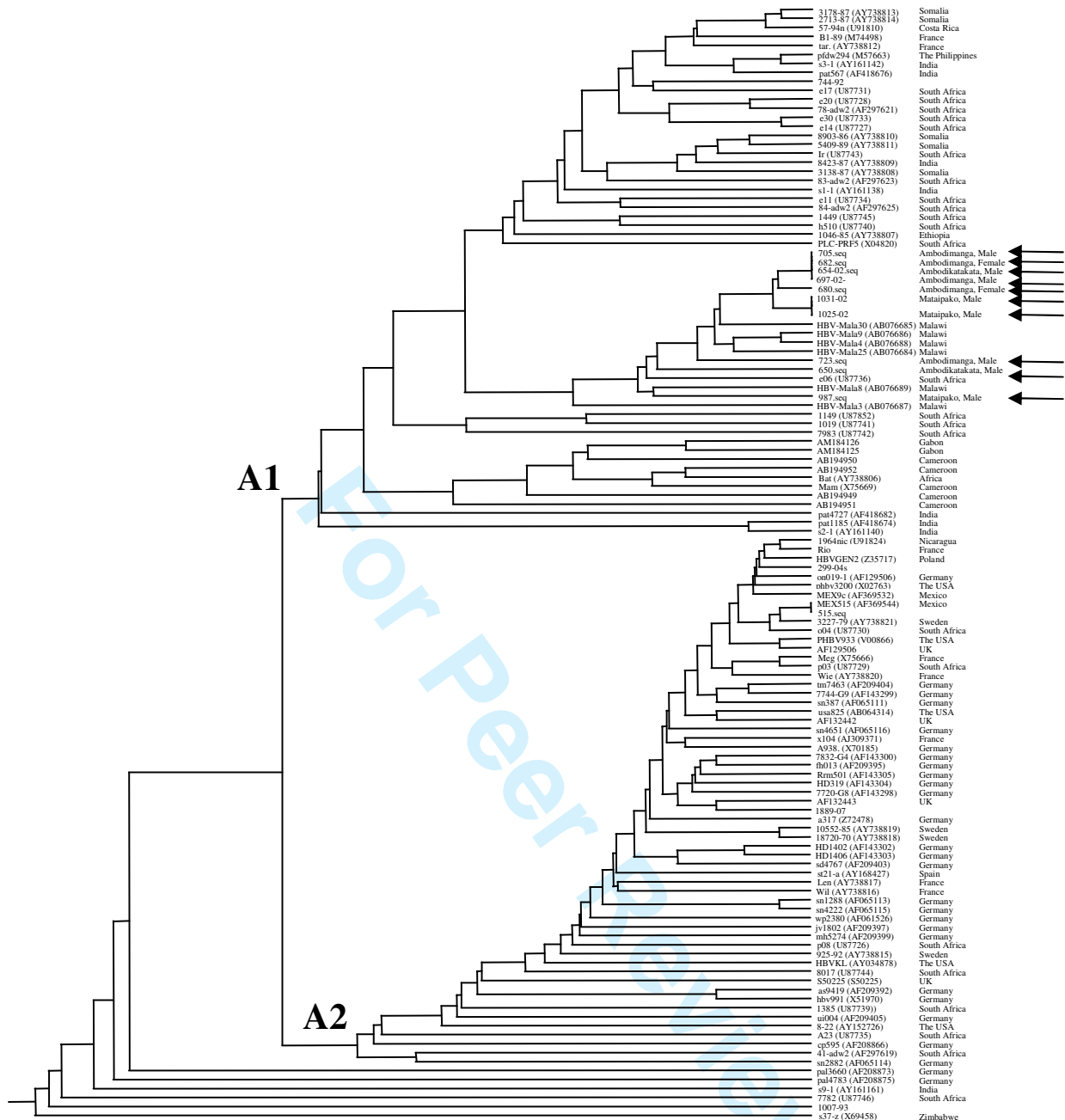
Figure 1

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.

For Peer Review

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

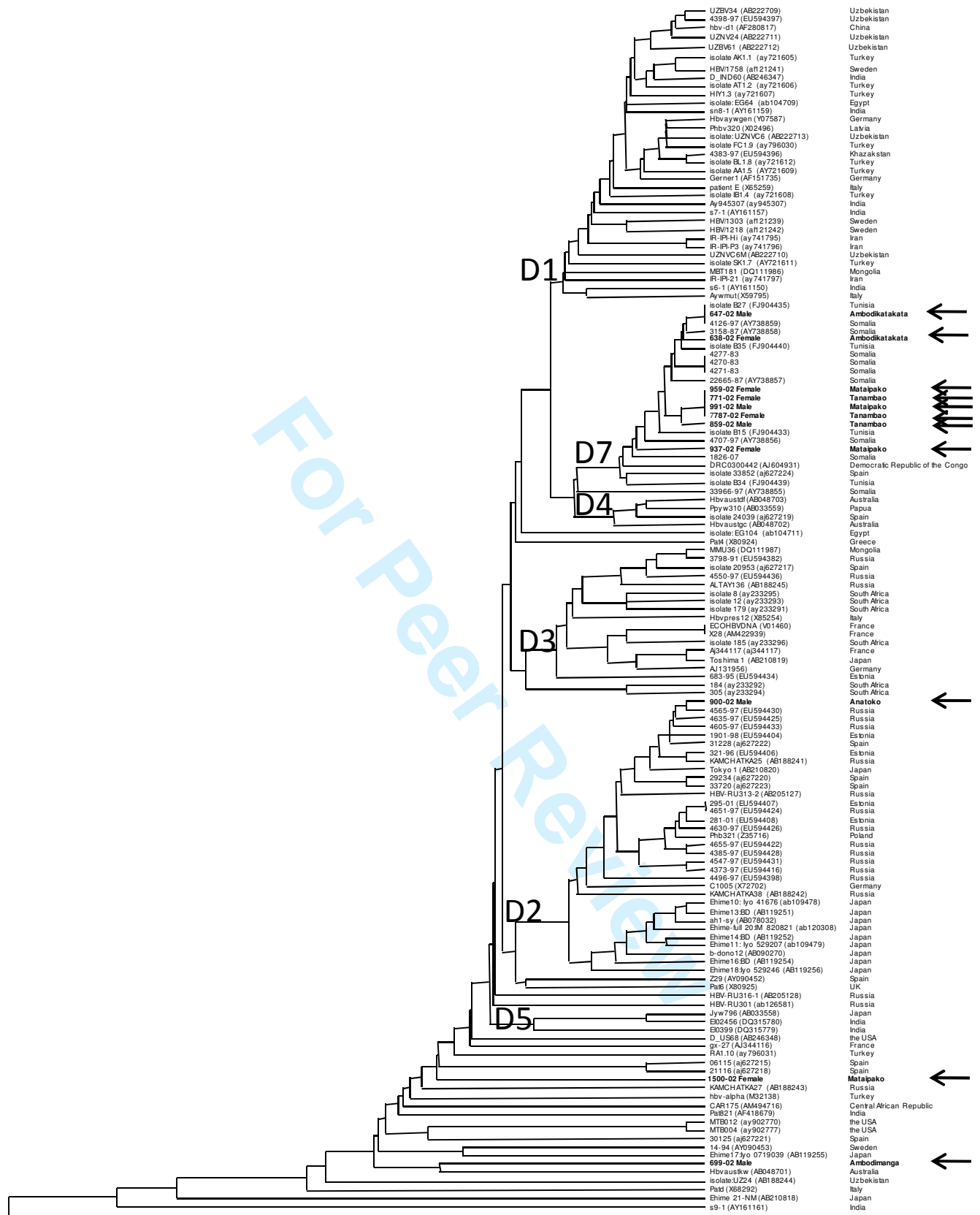


A1

A2

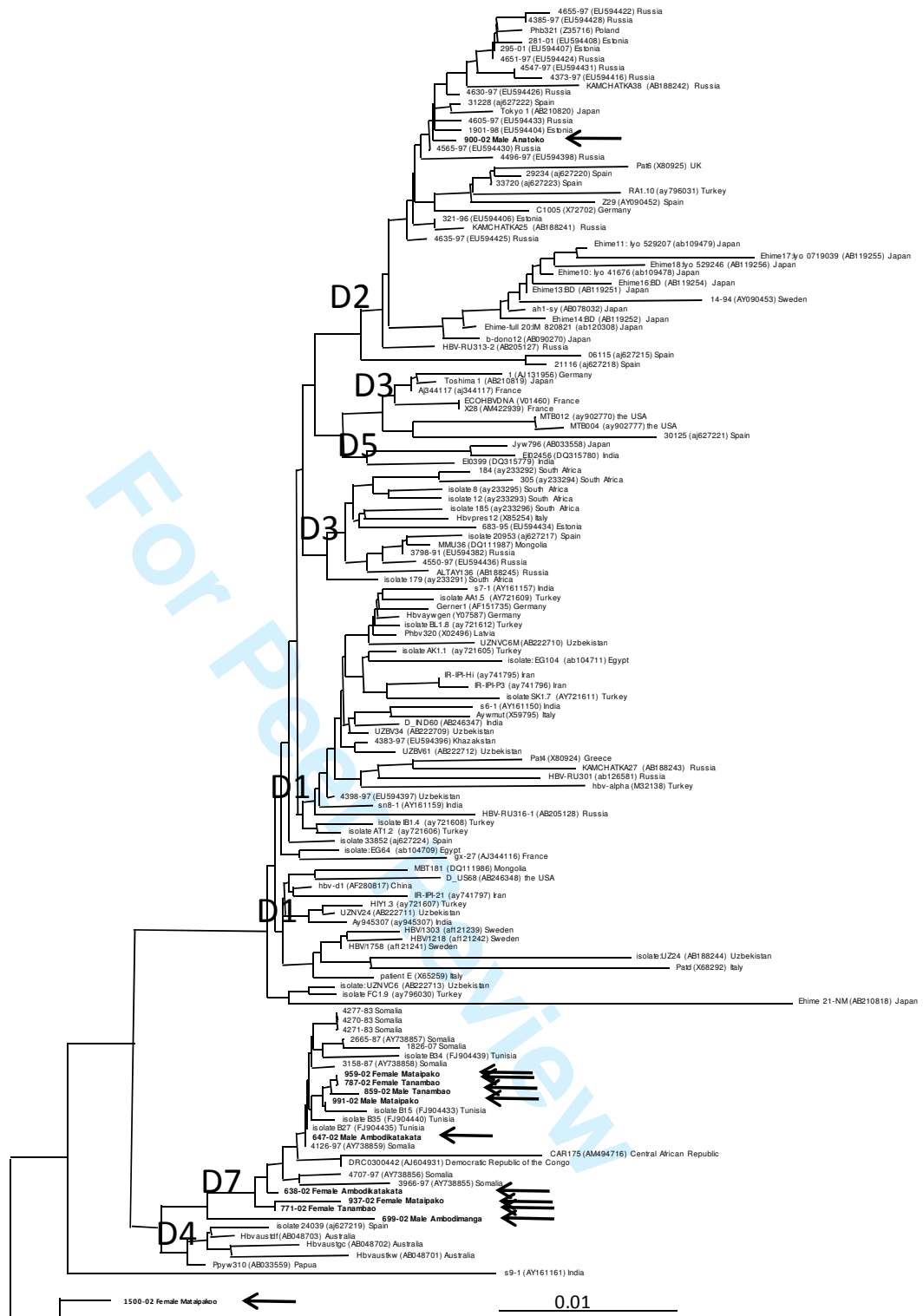
0.01

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60



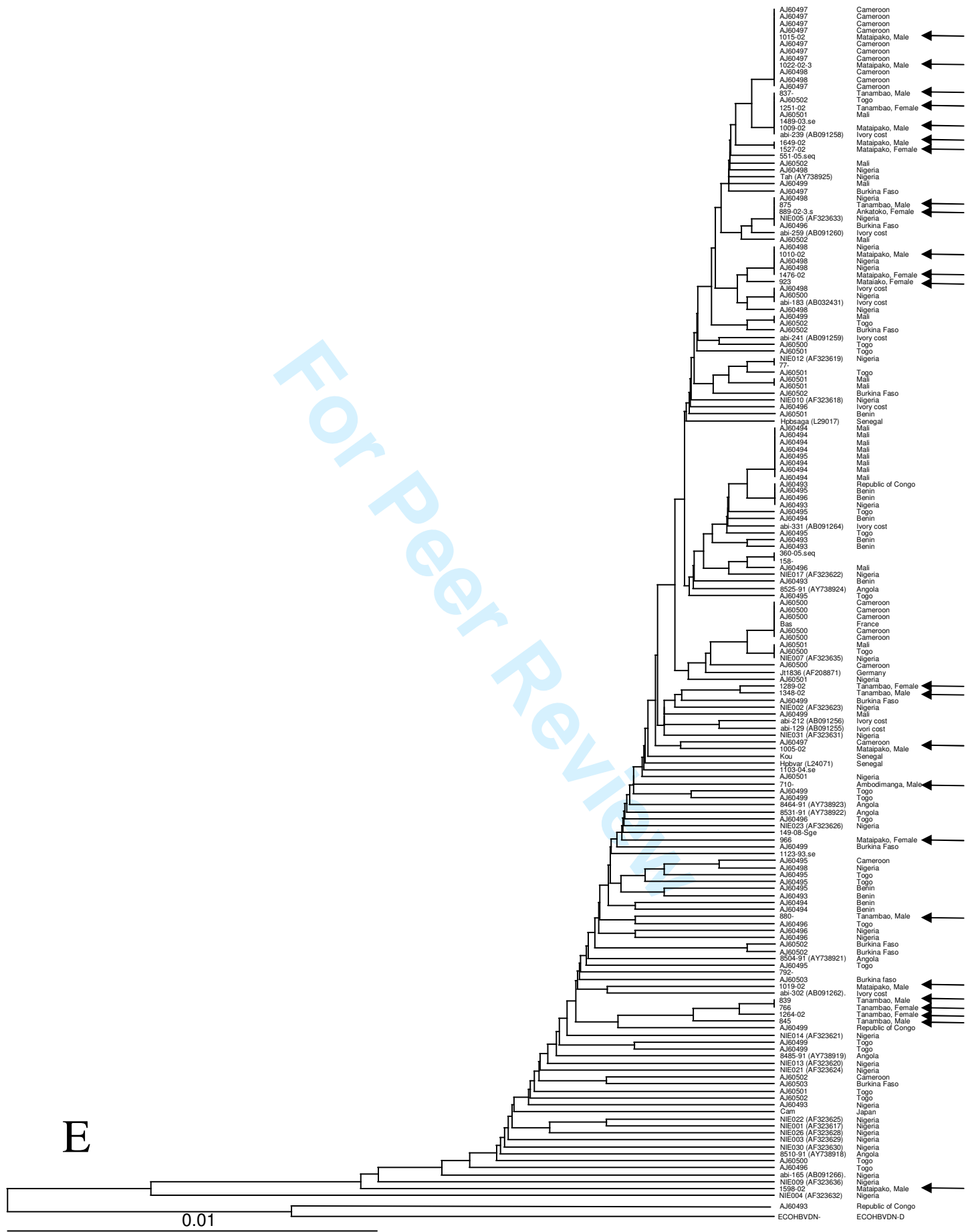
Genotype E

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60



Genotype E

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60



E

0.01