NEW MICROBIOLOGICA, **31**, 189-194, 2008

Restriction fragment length polymorphism analysis and direct sequencing for determination of HBV genotypes in a Turkish population

Ruchan Yazan Sertoz[1], Selda Erensoy[1], Suzan Pas[2], Tijen Ozacar[1], Hubert Niesters[2]

[1]Ege University Medical Faculty Department of Microbiology and Clinical Microbiology, Izmir, Turkey;
[2]Erasmus MC, University Medical Center Rotterdam, Department of Virology

SUMMARY

The aim of this study was to analyze the restriction fragment length polymorphism and direct sequencing results in genotyping of hepatitis B virus from a Turkish population in a clinical virology laboratory.
Serum samples of 54 chronic hepatitis B patients attending the Ege University Hospital were studied. Sequences of partial S gene PCR products were analysed and RFLP was performed. Fifty-three isolates could be identified by direct sequencing as genotype D. One sample needed to be cloned and determined as genotype D. Forty-two isolates were genotyped as D with RFLP according to published determinative patterns. Twelve isolates had undefined patterns. Eight of them suggested a mixture of isolates with different patterns and cloning of these samples confirmed the presence of heterogeneous isolates. Four isolates with undefined pattern were determined as genotype D by direct sequencing. All the studied isolates were genotype D.
The results of this studied population suggest that RFLP is suitable for HBV genotyping in a routine clinical virology laboratory setting. However sequence analysis and even cloning may be needed to clarify indeterminate results.

KEY WORDS: HBV, Genotyping, Sequencing, RFLP

Received July 04, 2007 *Accepted* December 27, 2007

INTRODUCTION

Despite the availability of HBsAg vaccines and mass immunization schemes, hepatitis B virus (HBV) infection continues to be a major public health problem. There are over 400 million chronic carriers worldwide and up to 40% patients may develop cirrhosis and hepatocellular carcinoma (Yuen *et al.,* 2002). Genetic classification based on the comparison of complete HBV genomes has defined eight genotypes, named A-H (EASL jury., 2003, Arauz *et al.,* 2002). Viral genomes have more than an 8% variation among different genotypes (Alestig *et al.,* 2001). Genotyping of HBV became important and clinical interest has increased because of the differences in pathogenesis, response to therapy and progression to hepatocellular carcinoma. In Turkey, nearly four million people are infected with HBV (Tasyaran, 2003) and studies have shown that genotype D is dominant (Yalcin *et al.,* 2004). Norder (Norder *et al.,* 2004) showed genetic diversity of the subgenotypes and their geographical distribution and concluded that subgenotype information may help to complement genetic data.
Several methods have been used for HBV genotyping including direct sequencing, RFLP (restriction fragment length polymorphism) and line probe assay (Lindh *et al.,* 1998, Naito *et al.,* 2001, Lindh *et al.,* 1997, Mbayed *et al.,* 1998). The accepted gold standard for genotyping of HBV is phylogenetic analysis. However, RFLP

Corresponding author
Ruchan Yazan Sertoz
Ege University Medical Faculty Microbiology Department
Izmir, Turkey
E-mail: ruchan.sertoz@ege.edu.tr

is a simple, suitable and cheap method for clinical virology laboratories. The aim of this study was to evaluate the suitability of RFLP in a routine clinical setting and analyze the RFLP patterns and direct sequencing results in genotyping of HBV in a clinical virology laboratory.

MATERIAL AND METHODS

Samples
Serum samples from 54 patients with chronic hepatitis B sent to the clinical virology laboratory of Ege University Hospital for HBV genotyping were included in the study. Patients were informed by their doctors and agreed for HBV genotyping to be performed. Thirty-two of the samples were HBeAg positive, 22 were HBeAg negative.

Extraction of HBV DNA from sera
Nucleic acid isolation from serum samples was done by High Pure Viral Nucleic Acid Kit (Roche Applied Science, USA).

HBV DNA PCR and genotyping
A selected genomic region of HBV gene was amplified by PCR with primers P7 and P8 (nt positions 256-278 and 796-776; according to unique EcoRI site) (Lindh *et al.,* 1997). Sequence reaction and analyzes of sequencing products were performed with HBS1 primer (nt 455-474) and ABI Prism 310 Genetic Analyzer (Perkin Elmer, USA) (Mbayed *et al.*, 1998).
DNA sequences were aligned, phylogenetic trees were constructed with HBV representative sequences from the literature and GenBank. Genotypes were determined according to groupings with HBV sequences of different genotypes and subgenotypes from GenBank. Nucleotide sequences were converted to amino acid sequences to determine HBsAg subtypes by alignment and comparison with the reported isolates. All these procedures were performed with DNAStar software (DNAStar INC, USA).
Phylogenetic trees were constructed by unweighed pair grouping method with arithmetic mean (UPGMA) and neighbor joining (NJ) methods with Treecon version 1b.
UPGMA (genetic distance and algorithmic clustering) method is preferred when the rate of nucleotide substitution per site per year is constant. NJ (genetic distance, clustering, algorithm) method is chosen when the nucleotide evolution is not constant especially for the regions with a high mutation rate. Distance estimation method Jukes Cantor assumes that the nucleotide variation rate for every nucleotide is the same. HBV is a highly mutated DNA virus but the mutation rate for genotypes is accepted as constant.

RFLP
PCR products were digested with Tsp5091 (AATT) and HinfI (GANTC) restriction enzymes. Restriction patterns were defined according to Lindh *et al.* (1997).

Cloning
PCR products with indeterminate RFLP patterns and electropherogram were amplified with the primers YMDD1 (nt371-397) and YMDD2 (nt839-861) and then were cloned into the pCR 2.1 TA vector (InVitrogen, Leek, The Netherlands). The clones were transformed to One Shot cells (INVαF'- InVitrogen, Leek, The Netherlands) and then inoculated to LB agar. PCR reactions were performed directly on recombinant colonies. Sequences of the strains that were cloned were compared with the other strains using the DNAStar software.

RESULTS

RFLP
Forty-two out of 54 samples (77.8%) could be genotyped according to the definitions of Lindh (Lindh *et al.,* 1997). Forty isolates had T8+H1 and two had T8+H3 patterns which were classified as genotype D.
Patterns of twelve isolates were not defined by Lindh. Eight were thought to be mixture of T8 and T11 patterns; seven of these had [T8+T11]+H1 and one had [T8+T11]+H3. TSP5091 restriction at nt480 produces fragments of 109bp and 173 bp (T8 pattern). A fragment of 282bp is observed if it cannot restrict at nt480 (T11 pattern). Heterogeneous sequences both with and without TSP5091 restriction at nt480 produce 109bp, 173bp fragments together with 282bp fragments. Four isolates had an undefined [T11+H1] pattern. The distrubution of RFLP patterns are shown in Table 1. Examples of TSP5091 and HinfI diges-

TABLE 1

Tsp5091 nt restriction site				*HinfI nt restriction site*					
480	*589*	*633*	*Tsp5091 pattern*	*481*	*529*	*HinfI pattern*	*Combined pattern*	*Genotype***	*No. of sequences*
+	+	+	T8	-	+	H1	T8+H1	D	40
+	+	+	T8	-	-	H3	T8+H3	D	2
(+/-)*	+	+	T8+T11	-	+	H1	T8+T11+H1	undefined (D)	7
(+/-)*	+	+	T8+T11	-	-	H3	T8+T11+H3	undefined (D)	1
-	+	+	T11	-	+	H1	T11+H1	undefined (D)	4
Total number of samples									54

Legend: The distrubution of samples according to RFLP patterns. *(+/-) Heterogeneous sequences both with and without Tsp5091 restriction site at nt480. **Definition of Lindh et al., 1998. The other patterns were not found in this definition [T8+T11]+H1; [T8+T11]+H3; T11+H1. It is confirmed that all these isolates are genotype D by sequencing.

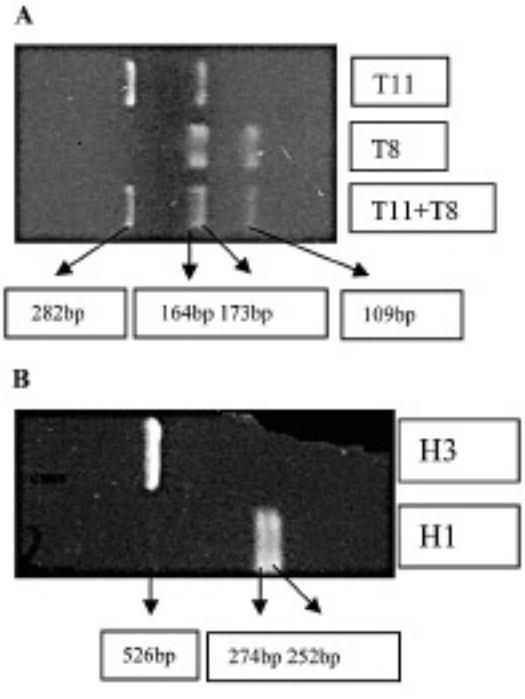
tion patterns are shown in Figure 1A and 1B. In seven mixed patterns three were from HBeAg positive and four were from HBeAg negative patients. HBeAg status had no significant relation to the patterns or groupings (data not shown).

Direct sequence analysis and cloning

Fifty-three samples (98.1%) were genotyped as D with direct sequencing and phylogenetic tree analysis. One sample with [T8+T11]+H1 mixed RFLP pattern could not be genotyped with direct sequencing. Cloning of this sample showed two minor variants among seven independent clones. One with nucleotide deletions between nucleotides 476 and 515 and another with a substitution of C instead of T (AATT® AACT) at nucleotide 483 were observed. These results show the presence of mixed sequences with TSP5091 restriction site (AATT) and without this site (-TT and AACT). Sequences with deletions between 476 and 515 and sequences of [T11+H1] RFLP pattern were not found among the submitted sequences in GenBank.

Direct sequencing analysis of seven sequences with a mixture of [T8+T11] RFLP patterns showed double peaks at nucleotides 482 or 483, suggesting the mixture of viral sequences with and without TSP5091 restriction site. Cloning of these samples confirmed the presence of heterogeneous isolates.

Similarity analysis of the sequences showed 92-100% similarity between the selected genotype D sequences and our samples (DNAStar INC, USA). All our isolates grouped as genotype D had the representative sequences shown Figure 2.



Legend: A. *TSP5091* digestion RFLP pattern
B. *Hinf1* digestion RFLP pattern

FIGURE 1

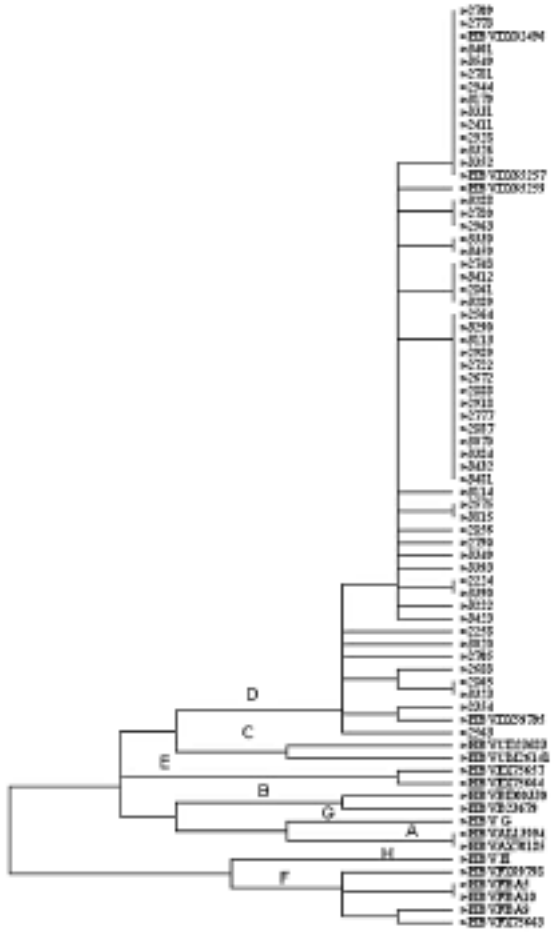
Analysis of subgenotypes

We could not obtain significant clusters for genotyping with phylogenetic tree analysis (both with NJ and UPGMA), probably due to shorter sequences compared to those used by Norder.

We investigated nucleotide differences to discriminate the subgenotypes according to Norder (Norder *et al.,* 2004). Discriminative differences from D1 in significantly clustered subgenotypes were noted and shown in Figure 3. These discriminative differences were: G at nt506, A at 533 for D2; T at 190, C at 493, C at 518, T at 528 together with A at 533 for D3; A at 493, C at 514, A at 574, C at 592, T at 668 for D4. According to discriminative nucleotide differences, 51 of our isolates were determined as D1, while three were D2 (Figure 3).

HBsAg subtypes

All 54 isolates were determined as ayw according to sequences converted to HBsAg aminoacids. Fifty of them were ayw2, two were ayw3. Two could be determined as ayw, but could not be subtyped further. Identical sequences were found in GenBank identified as ayw.



Legend: The phylogenetic tree of the HBV DNA sequences between nt.502-774 (EcoR1 site) was constructed by Treecon version 1b (Jukes Cantor distance estimation and UPGMA). Mutations related to lamivudine resistance are avoided by the tree analysis. All of our samples are clustered with selected HBV genotype D sequences from GenBank. Our sequences are the ones with four digit numbers.

FIGURE 2

nt190	nt493	nt506	nt514	nt518	nt528	nt533	nt574	nt592	nt668	
T	G	C	T	T	A A	C	A C G A C T C C		A T T	C C T G X02496-D1
.	.	.	.	.	.	.	. A		.	. . . . 18-94-D1
.	.	.	.	.	.	.	. A		.	. . . . 15-94-D1
.	.	.	.	.	.	.	. . . . .		.	. . . . 1331-93-D1
.	.	.	.	.	.	.	. A		.	. . . . 1281-93-D1
.	.	.	.	.	.	.	. . . . .		.	. . . . 7794-85-D1
.	.	.	.	.	.	.	. . . . .		.	. . . . AF209391-D1
.	.	.	.	.	.	.	. . . . .		.	. . . . AY161159-D1
.	.	.	.	.	.	.	. . . . .		.	. . . . 161150-D1
.	.	.	.	.	.	.	. . . . .		.	. . . . X59795-D1
.	.	.	.	.	.	.	. . . . .		.	. . . . X65259-D1
.	.	.	.	.	.	.	. . . . .		.	. . . . M32138-TURKEY
.	.	.	.	.	.	.	. . . . .		.	. . . . AF151735-D1
.	.	.	.	.	.	.	. . . . .		.	. . . . 1616-99-D1
.	.	.	.	.	.	.	. . . . .		.	. . . . AY161157-D1
.	.	.	.	.	.	.	. . . . .		.	. . . . AF280817-D1
.	.	.	.	.	.	.	. . . . .		.	. . . . AF061528-D1
.	.	.	.	.	.	.	. . . . .		.	. . . . AF209398-D1
.	.	.	.	.	.	.	. A		.	. . . . 304-94-D1
.	.	.	.	.	.	.	. . . . .		.	. . . . 8208-87-D1
-	-	-	-	.	.	.	. . . . .		.	 2224
-	-	-	-	.	.	.	. . . . .		.	 2258
-	-	-	-	.	.	.	. . . . .		.	 2354
-	-	-	-	.	.	.	. . . . .		.	 2411
-	-	-	-	.	.	.	. . . . .		.	 2563
-	-	-	G	.	.	.	. . . . . A .		.	 3323
-	-	-	G	.	.	.	. . . . . A .		.	 2688
-	-	-	G	.	.	.	. . . . . A .		.	 2865
.	.	.	G	.	.	.	. . . . . A .		.	. . . . Z35716-D2
.	.	.	G	.	.	.	. . . . . A .		.	. . . . AB090268-D2
.	.	.	G	.	.	.	. . . . . A .		.	. . . . 9842-85-D2
.	.	.	G	.	.	.	. . . . . A .		.	. . . . 8950-90-D2
.	.	C	G	.	.	.	. . . . . A .		.	. . . . AB090269-D2
.	.	.	G	.	.	.	. . . . . A .		.	. . . . AB090270-D2
.	.	.	G	.	.	.	. . . . . A .		.	. . . . AF209401-D2
.	.	.	G	.	.	.	. . . . . A .		.	. . . . X72702-D2
.	.	.	G	.	.	.	. . . . . A .		.	. . . . 11456-92-D2
.	.	.	G	.	.	.	. . . . . A .		.	. . . . A1229-D2
.	.	G	G	.	.	.	. . . . . A .		.	. . . . AB078032-D2
.	.	.	G	.	.	.	. . . . . A .		.	. . . . AB078033-D2
.	T	C	.	.	.	.	C T . . . . A .		.	. . . . 1483-70-D3
.	T	C	.	.	.	.	C T . . . . A .		.	. . . . 1596-99-D3
.	T	C	.	.	.	.	C T . . . . A .		.	. . . . 921-92-D3
.	T	C	.	.	.	.	C T . . . . A .		.	. . . . 2373-93-D3
.	T	C	.	.	.	.	C T . . . . A .		.	. . . . 33234-73-D3
.	T	C	.	.	.	.	C T . . . . A .		.	. . . . AF061523-D3
.	T	C	.	.	.	.	C T . . . . A .		.	. . . . AF209396-D3
.	T	C	.	.	.	.	C T . . . . A .		.	. . . . V01460-D3
.	T	C	.	.	.	.	C T . . . A A .		.	. . . . X75668-D3
.	T	C	.	.	.	.	C T . . . . A .		.	. . . . 1823-02-D3
.	T	C	.	.	.	.	C T . . . . A .		.	. . . . 1272-00-D3
.	.	A	.	.	C	.	. . . . . A		.	C . T . . 2473N-D4
.	.	A	.	.	C	.	. . . . . A		.	C . T . . 3153N-D4
		A	.	.	C	.	. . . . . A		.	C . T . . AB033559-D4
		A	.	.	C	.	. . . . . A		.	C . T . . AB048701-D4
		A	.	.	C	.	. . . . . A		.	C . T . . AB048702-D4
		A	.	.	C	.	. . . . . A		.	C . T . . AB048703-D4
		A	.	.	C	.	. . . . . A		.	C . T . . X75662-D4

Legend: Boxshade alignment of the sequences from our isolates with the reported sequences as representative subgenotypes of D (Norder et al., 2004). Five D1 examples out of 51 of our samples and three D2 samples are also shown (2224,2258,2354,2411,2563 and 3323,2688,2865) nt 190 and nt493 are not included in our PCR products.

FIGURE 3

DISCUSSION

HBV can be classified into eight genotypes A-H, according to current knowledge based on an intergroup divergence of 8% or more in the complete viral genome. Genotype A is the predominant genotype in northern Europe; genotypes B and C are predominant in East Asia and the Far East; genotype D is found worldwide, but prevails in the Mediterranean area; genotype E is prevalent in western-Subsaharan areas and genotype F is likely to be present in populations with origins on the American continent (Arauz *et al.,* 2002, Yalcin *et al.,* 2004, Mbayed *et al.,* 1998).

It seems that different genotypes might have different pathogenesis. According to Lee *et al.,* genotype C carriers were more often positive for HBeAg than genotype B carriers (Lee *et al.,* 2003). There are also papers from India where genotypes A and D are prevalent, showing that genotype D is associated with more severe disease and may predict the occurrence of hepatocellular carcinoma in younger patients (Clarke *et al.,* 2002). The clinical relevance of HBV genotyping needs to be investigated and evaluated. Genotyping may be beneficial in the management of HBV infection since there is a relationship between genotype and the outcome of infection. Analysis of subgenotypes may add further explanations for different clinical outcomes and response to antiviral therapy.

There are few studies on HBV genotyping in Turkey. All the researchers found genotype D in different patient populations as in this study (Kulah *et al.,* 2002, Misirlioglu *et al.,* 2002, Bozdayi *et al.,* 2001, Yazan Sertoz *et al.,* 2003). Genotyping by sequencing remains the gold standard when performed on the correct genome regions and when adequate models of evolution and methods are used. However, it is a labor intensive and expensive method (Clarke *et al.,* 2002). Routine clinical laboratories need simple and economic reliable methods suitable to laboratory conditions and sample population. These methods need to be evaluated for their efficiency in laboratory settings.

RFLP, which is a simple classical molecular biology technique may be efficient in genotyping when the differences at restriction sites are significant between different genotypes.

HBV subgenotypes have been introduced with intergroup divergences less than 4% throughout the complete genome. The clinical relevance of subgenotypes is not known. Nine Turkish HBV sequences from GenBank were determined as D1 by Norder (Norder *et al.*, 2004). Our sequences were shorter than the analyzed sequences. There was no problem in determining the genotypes by phylogenetic tree with this length of sequences (nt502-774). However, defined subgenotypes were not precisely grouped in the dendrogram while longer sequences (nt155-835) revealed correct groupings of subgenotypes (data not shown). Norder *et al.* also report that bootstrap values did not support subgenotype groupings at the S gene level, whereas there was a strong concordance in most of the subgenotypes according to complete genome phylogenetic analysis with significant bootstrap support.

We could not subgenotype by phylogenetic tree with short sequences. In order to have subgenotype data, we analyzed aligned HBV genotype D sequences of different subgenotypes to distinguish nucleotide differences and noted the ones belonging to the significantly clustered subgenotypes as defined in results. Two of these which distinguish D3 were not covered by our sequences (nt 190 and 493). It was noted that D3 sequences were the major problem in the dendrograms with the shortened length of our sequences. D3 had two more specific differences (nt518, 528; which are not detected in our isolates) in the region of our studied PCR products. According to these differential nucleotides, 51 (94.4%) of our isolates were classified as D1 and three were D2. This finding is concordant with the nine D1 sequences reported previously (Norder *et al.,* 2004).

There is a need for discussion and then a consensus on HBV genotyping in Turkey since it is shown that genotype D is dominant. If the efficiency of genotyping is accepted, a simple and a cheap reliable method such as RFLP may be preferred.

In conclusion, the results of this studied population suggest that RFLP is suitable for HBV genotyping in a routine clinical virology laboratory setting. In different settings when more than one genotype circulate some additional limitations may arise, due to a more heterogeneous RFLP pattern or the limitations caused by the genomic region chosen. Heterogenous sequences of sample populations may cause problems in geno-

typing by RFLP or direct sequencing. In these cases both methods or supplemental tests, even cloning may be necessary (Lindh *et al.,* 1997). This conclusion applies to a HBV genotype D population.

ACKNOWLEDGEMENTS
We thank to Seyhan Dargi, Yüksel Akba for their excellent technical assistance.

REFERENCES

ALESTIG E., HANNOUN C., HORAL P., LINDH M. (2001). Hepatitis B virus genotypes in Mongols and Australian Aborigines. *Arch. Virol.* **146**, 2321-2329.

ARAUZ-RUIZ P., NORDER H., ROBERTSON B.H., MAGNIUS L.O. (2002). Genotype H: a new Amerindian genotype of hepatitis B virus revealed in Central America. *J. Gen. Virol.* **83**, 2059-2083.

BOZDAYI A.M., BOZKAYA H., TÜRKYILMAZ A.R., SARYODLU M., ÇETINKAYA H., ET AL. (2001). Nucleotide divergences in the core promoter and precore region of genotype D hepatitis B virus in patients with persistently elevated or normal ALT levels. *J. Clin. Virol.* **21**, 91-101.

CLARKE B., BLOOR S. (2002). Molecular genotyping of hepatitis B virus. *J. Clin. Virol.* **25**, S41-45.

EASL JURY. (2003). EASL International Consensus Conference on Hepatitis B. *J. Hepatol.* **39**, S3-S25.

KÜLAH C., YALINAY-ÇIRAK, M. (2002). Genotyping of turkish hepatits B virus isolates based on the sequencing of the S regions [abstract]. 2nd molecular and diagnostic microbiology congress proceedings and programme book, Ankara Microbiology Society April **60**, 21-25.

LEE C.M., CHEN C.H., LU S.N., TUNG H.D., CHOU W.J., WANG J.H., ET AL. (2003). Prevalence and clinical implications of hepatitis B virus genotypes in Southern Taiwan. *Scand. J. Gastroenterol.* **40**, 95-101.

LINDH M., ANDERSSON A.S., GUSDAL G. (1997). Genotypes, nt 1858 variants, and geographic origin of hepatitis B virus: large-scale analysis using a new genotyping method. *J. Infect. Dis.* **175**, 1285-1293.

LINDH M., GONZALEZ J.E., NORKRANS G., HORAL P. (1998). Genotyping of hepatitis B virus by restriction pattern analysis of a pre-S amplicon. *J. Virol. Methods*. **72,** 163-174.

MBAYED V.A., LOPEZ J.L., TELENTA P.F.S., PALACIOS G., BADIO I., FERRO A., ET AL. (1998). Distrubution of hepatitis B virus genotypes in two different pediatric populations from Argentina. *J. Clin. Microbiol.* **36**, 3362-3365.

MISIRLIO LU, M., KAYIN E., AKMAN E., TUNCER S. (2002). Hepatitis B virus genotypes and viral load analysis in pre and postvaccination sera from carrier children [abstract]. 2nd Molecular and diagnostic microbiology congress proceedings and programme book, Ankara Microbiology Society April 21-25, 63.

NAITO H., HAYASHI S., ABE K. (2001). Rapid and specific genotyping for hepatitis B virus corresponding to six major genotypes by PCR using type-specific primers. *J. Clin. Microbiol.* **39**, 362-364.

NORDER H., COUROUCE A.M., COURSAGET P., ET AL. (2004). Genetic diversity of hepatitis B virus strains derived worldwide: genotypes, subgenotypes, and HBsAg subtypes. *Intervirology.* **47**, 289-309.

TAŞYARAN M. (2003). HBV İnfeksiyonu epidemiyoloji. In: Viral Hepatit (Tekeli E, Balik İ, editors.) 1st ed p: 121-129.

YALCIN K., DEGERTEKIN H., BAHCECIOGLU I.H., ET AL. (2004). Hepatitis B virus genotype D prevails in patients with persistently elevated or normal ALT levels in Turkey. *Infection.* **32**, 24-29.

YAZAN-SERTÖZ R., ÖZACAR T., ERENSOY S. (2003). Genotyping of Hepatitis B virus by direct sequence analysis of PCR products [abstract]. *J. Clin. Virol*. **27**, Suppl 1: 53.

YUEN M.F., YUAN H.J., WONG K.H., ET AL. (2002). A large population study of spontaneous HBeAg seroconversion and acute exacerbation of chronic hepatitis B infection: implications for antiviral therapy. *Gut*. **52**, 416-419.