

Hepatitis B Virus Genotypes in Latvia

Irina Sominskaya¹, Maria Mihailova¹, Juris Jansons^{*1}, Diana Legzdina¹, Gunita Sudmale¹, Paul Pumpens¹, Frida Arsha², Irena Davidjuka², Jazeps Keish², Valentina Sondore², Baiba Rozentale² and Ludmila Viksna²

¹Latvian Biomedical Research and Study Centre, Ratsupites 1, Riga LV-1067, Latvia

²Infectology Center of Latvia, Linezera 3, Riga LV-1006, Latvia

Abstract: The aim of this study was to investigate the properties of HBV genomes isolated from patients with acute and chronic hepatitis, liver cirrhosis, and hepatic coma admitted to the Infectology Center of Latvia. HBV genotypes and HBsAg subtypes were determined by direct S-gene sequencing. Genotypes D (72.17%) and A (27.96%) were predominant. Only one case (0.87%) of HBV genotype E was found. Prevalence of genotype D over genotype A was more strongly pronounced in isolates from patients with acute hepatitis than from patients with chronic hepatitis— 74%-24% and 68%-32%, respectively. Number of previously described and some new mutations (Ser21Thr in preC region, Tyr134Lys, Phe134Leu, Ile110Met, Thr118Ile and Ser136Tyr in S-gene) were discovered by direct sequencing of PCR fragments corresponding to preS/S and preC/C regions and to a fragment of the X gene of HBV genome. The most frequently mutations were found in the region of the basal core promoter, preC and major immunodominant region of core protein and domain *a* of S protein. In general the rate of mutations discovered in chronic patients was just 1,23 times higher than in patients with acute hepatitis, at the same time the ratio for the hot-spot preC 1896 stop mutation was higher and reached 1,81. A possible role of discovered mutations in the HBV pathogenesis is discussed. The findings are relevant to diagnosis and prognosis of HBV-induced liver disease.

Keywords: Hepatitis B virus, genotype, subtype, variant, mutation.

INTRODUCTION

The clinical course of hepatitis B virus (HBV) infection varies from self-limited and asymptomatic acute to fulminant hepatitis. A definite population of patients with acute HBV infection (AH) develops chronic liver disease: chronic hepatitis (CH), liver cirrhosis (LC), and hepatocellular carcinoma (HCC). HBV is classified into 8 genotypes from A to H based on more than 8% sequence divergence of the entire genome sequences [1-3]. The HBV genotypes appear to be associated with particular geographic distribution, ethnicity, and possibly diagnoses and clinical outcome of the liver disease. At first, the HBV genotypes have distinct geographical distributions, e.g. A – Northwest Europe, North America, and Central Africa; D – Mediterranean area, Middle East, and India; E – sub-Saharan Africa. The HBV genotypes are divided into clusters or subgenotypes with 4 to 8% sequence divergence. Existence of two subgenotypes within the genotype A is described for South Africa [4]. These subgenotypes are assigned as A1 and A2 and the major difference between them is associated with to amino acid (aa) residues 207 and 209 in the HBs protein S (Asn/Leu and Ser/Val substitutions for A1 and A2, respectively). It has been pointed out that European HBV strains belong to the subgenotype A2, whereas Afro-Asian strains belong to the subgenotype A1. A subgenotype A3, found in Cameroon,

is a recombinant of the genotypes A and E [5]. The most widespread in the world is the genotype D. This genotype is divided into four subgenotypes D1 to D4. The geographic distribution of these subgenotypes is less restricted than that for the genotype A, although the strains from the Middle East mainly belong to D1, from South Africa and Alaska to D3, and from Oceania and Somalia to D4 [6]. None of these subgenotypes is characterized with specific aa substitutions.

Traditionally, HBV is classified into 9 hepatitis B surface antigen (HBsAg) subtypes: *adw2*, *adw4*, *adrq⁺*, *adrq⁻*, *ayw1*, *ayw2*, *ayw3*, *ayw4* and *ayr* according to the phenotypic expression of the HBs-encoding part of the gene S of different HBV genotypes [7]. Epidemiological studies show that the prevalence of these subtypes also varies in different parts of the world and that they probably have a different clinical significance [8,9]. The *adw2* and *ayw1* subtypes usually belong to the genotype A, whereas the *ayw2*, *ayw3*, *ayw4* and *adw3* subtypes belong to the genotype D. The *ayw4* subtype appears in the genotype E.

Despite the introduction of anti-HBV vaccination in 1997 and constant improvement of medical care in Latvia, the number of hepatitis B cases remains high. In previous papers we described appearance of HBV genotypes in Latvia [10,11] and their distribution in two Latvian clinics: the Kidney transplantation Centre and the Hematology Unit of the Children's Clinical University Hospital in Riga [12,13]. Due to the specific nature of the clinics obtained data represented exclusively the HBV genome distribution for different sorts of chronic HBV infection.

*Address correspondence to this author at the Latvian Biomedical Research and Study Centre, Ratsupites 1, Riga, LV-1067, Latvia; Tel: 371-7808212; Fax: 371-7442407; E-mail jansons@biomed.lu.lv

Table 1. Distribution of Patients Admitted to Study by Sex, Age, Gender and Major HBV and HCV Markers

Desase	Total	Sex	Number	Age	HBV DNA +	HBV DNA -	a-HCV +
Acute hepatitis	89	m	47	34.82±17.42	42	5	8
		f	42	44.22±18.26	40	2	3
Chronic hepatitis	33	m	24	40.91±17.65	19	5	7
		f	9	36.00±19.87	6	3	0
Other liver damages	9	m	6	27.67±2.89	5	1	1
		f	3	69.50±7.78	3	0	0
Total	131	m	77	35.69±17.08	66	11	16
		f	54	44.25±18.92	49	5	3

The aim of this study is to evaluate a present epidemiological distribution of the HBV genotypes, subgenotypes, and subtypes among patients of the Infectology Center of Latvia (ICL), a major medical institution in this field in Latvia, whose main specialization is acute HBV infection. We present here the appearance of mutations within the HBV genome regions EnhII/BCP/X-preC/C and preS/S against the HBV genome prototypes. Obtained results shows that the distribution of the HBV genotypes in Latvia remains the same for all kinds of HBV-induced liver disease patients and diagnoses, and the profile of the medical institutions involved. More than 20-years long persistence of area-specific HBV strains with preserved "old" mutations reveals a clear predominance of "local" HBV genotypes, despite the rapid development and growing migration rate in Latvia.

MATERIALS AND METHODS

Patients

A total of 131 (54 female and 77 male) consecutive patients of the ICL from December 1st 1997 to December 31st 2006 were admitted to the study. Diagnosis: AH, CH, LC, autoimmune hepatitis, and hepatic coma (Table 1). Diagnosis was based on traditional clinical, biochemical, and immunochemical hepatitis B marker tests. The mean age of the female and male patients was 44.25 ± 18.92 years and

35.69 ± 17.08, respectively. State agency "Infectology Center of Latvia" and the Scientific Council of the Biomedical Research and Study Centre approved the protocol of this study and appropriate research permission was obtained from the independent Ethics Committee of Latvia. Informed consent to participate into this study was obtained from all included patients.

Serological Tests

The serum samples were stored at -20 °C until assay was performed. All serum samples were tested for the presence of HBsAg (ETI-MAK-4, DiaSorin, Saluggia, Italy and Enzygnost HBsAg 5.0, Dade Behring, Marburg, Germany), total antibodies against hepatitis B core antigen (anti-HBc) (ETI-AB-COREK-2, DiaSorin, Saluggia, Italy), as well as for anti-HBc IgM, and anti-HBs. HBsAg positive samples were tested for hepatitis B e antigen (HBe) (ETI-EBK-2, DiaSorin, Saluggia, Italy) and antibodies against it (anti-HBe) (ETI-AB-EBK, DiaSorin, Saluggia, Italy). The study protocol was in agreement with rules of Latvian Ethics Committee.

PCR Amplification and Sequence Analysis

HBV DNA was extracted from 50 µl of serum with a commercially available DNA-RNA isolation kit based on

Table 2. PCR Primers Used in Study

Region of Analysis		Primer	Sequence
pres/S	2750-813 nt	17p	5' TTATTTACATACTCTTTGGAAGGC 3'
		8.15	5' AATGTATACCCAAAGACAG 3'
gene S (nested)	215-710 nt	S1p	5' TTGTTGACAAGAATCCTCACAATACC 3'
		S2p	5' GCCCTACGAACCACTGAACAAATGG 3'
preC/C	1741-2516 nt	M3	5' CTGGGAGGAGTTGGGGGA 3'
		p19	5' AGGTACTGTAGAGGAATAAAGCCC 3'
geneC (nested)	2045-2387nt	C1"	5' GTTCACCTCACCATACTGCACTCAGGC 3'
		C2'	5' GAGTCTTCTTCTAGGGACCTGCCTCG 3'
BCP/preC (nested)	1741-2114 nt	M3	5' CTGGGAGGAGTTGGGGGA 3'
		2.364	5' CCCAGGTAGCTAGAGTCAT 3'
gene X	1505-1825 nt	21p	5' CGTTCCAGCCGACCACGGGGCGC 3'
		2-Sp	5' AAAAAGTTGCATGGTGCTGG 3'

phenol/chloroform extraction (LyteH, Moscow, Russia). HBV genome fragments carrying the whole preS/S, preC/C regions, and a fragment of the X gene, which included the EnII/BCP region, were PCR-amplified with several primer sets (Table 2), as previously described [13]. When amplification of the whole fragments failed, we used nested PCR with inner primers. Amplified HBV DNA fragments were sequenced in both directions directly using the ABI Prism Dye Terminator Cycle Sequencing Ready Reaction (Applied Biosystems, Foster City, Calif., USA) and electrophoregrams were obtained on an ABI Prism 377 sequencer (Applied Biosystems). In all cases PCR primers were used as sequencing primers. For sequencing of the complementary strand of the HBV preS region, an additional primer p14 – 5' CTGTAA CACGAGCAGGGGTCCTAG 3' was used [14]. The sequences were edited manually with the BioEdit Sequence Alignment Editor and subsequently aligned in the FASTA format (<http://ngfnblast.gbf.de/docs/fasta.html>). The phylogenetic tree was constructed by the MEGA program using the UPGMA algorithm.

RESULTS AND DISCUSSION

Sera of patients were divided into three groups: first group from 89 patients with AH, second one from 33 patients with CH, and third one from 9 patients with LC, autoimmune hepatitis, hepatic coma, and fulminant hepatitis (FH). The last group was named “other liver damage” (Table 1). All AH sera patients were HBsAg and anti-HBc IgM positive. Sera of 19 patients contained anti-HCV antibodies. One serum of an AH patient contained anti-HAV antibodies and one of the CH patients was positive for HAV. Serological

data concerning presence of HBsAg and HBe/anti-HBe patient status are listed in Table 3.

All samples were tested for HBV DNA by PCR; HBV DNA was detected in sera from 115 of the 131 patients (87.79 %), 16 samples were negative (12.21 %) (Table 1).

Phylogenetic Analysis of HBV Genotypes

The genotype of each HBV sequence was determined by comparison with sequences of known HBV genotypes A-H in the GenBank and EMBL nucleotide databases using the BLAST program. Phylogenetic tree was constructed for the HBV gene S (Fig. 1). Reference sequences, which are shown with their accession numbers, were obtained from the EMBL/GenBank database. Despite the multinational society existing in Latvia and increasing population mobility during recent years, we found only genotypes A (26.96% of cases) and D (72.17% of cases) and only one case of a “foreign” genotype E (0.87%) in the studied patients. Phylogenetic analysis clearly showed similarity between sequences of the genotype A, whereas the genotype D sequences were more divergent. All genotype A isolates clustered together with A2 subgenotype sequences from the GenBank database and contained typical aa residue Ser at position 207 and Val at position 209. This means that the genotype A strains possess a rather low level of divergence and are “newcomers” to Latvia. The genotype D isolates had a more pronounced level of divergence and fell into three main clusters. The major cluster consisted of 48 isolates, which shared some similarity with the strains from the D2 subgenotype [6]. From this cluster, some subclusters were distinguished. Two isolates showed similarity with the strains from the D1

Table 3. Frequency of Specific Immunological Markers in Two Groups of Patients: AH, CH and other Liver Damages

Disease	Total	HBsAg+	HBsAg+	HBsAg+	HBsAg-
		HBeAg+	a-HBe +		a-HBe+
Acute hepatitis	89	27	26	1	13
Chronic hepatitis	33	16	8	3	5
Other liver damages	9	3	2	3	1

Table 4. Distribution of HVB Genotypes, Subgenotypes and HBsAg Subtypes between LIC Isolates. Number of Isolates with Mutations is Shown in Brackets

Disease	Total	Sex	Number	Genotypes, Subgenotypes and HBsAg Subtypes									
				D						A		E	
				D1	D2		D3		D?	S gene PCR negative	A2	S gene PCR negative	ayw4
				ayw2	ayw3	ayw2	ayw2	ayw3			adw2		
Acute hepatitis	82	m	42(25)	0	22(13)		2(2)	1(1)	0	5(1)	10(8)	2	0
		f	40(23)	0	21(11)		3(1)		0	7(5)	6(4)	2(1)	1(1)
Chronic hepatitis	25	m	19(9)	1	1		1(1)		1(1)	8(4)	4(3)	3(2)	0
		f	6(4)	0	1	1	2(2)		0	1(1)	1(1)	0	0
Other liver damage	8	m	5(4)	0	2(2)			0	0	0	3(2)	0	0
		f	3(2)	1	1(1)			0	0	1(1)	0	0	0
Total	115	m	66(40)	1	25(15)		4(4)		1(1)	13(4)	17(13)	5(2)	0
		f	49(29)	1	24(12)		5(3)		0	9(7)	7(5)	2(1)	1(1)

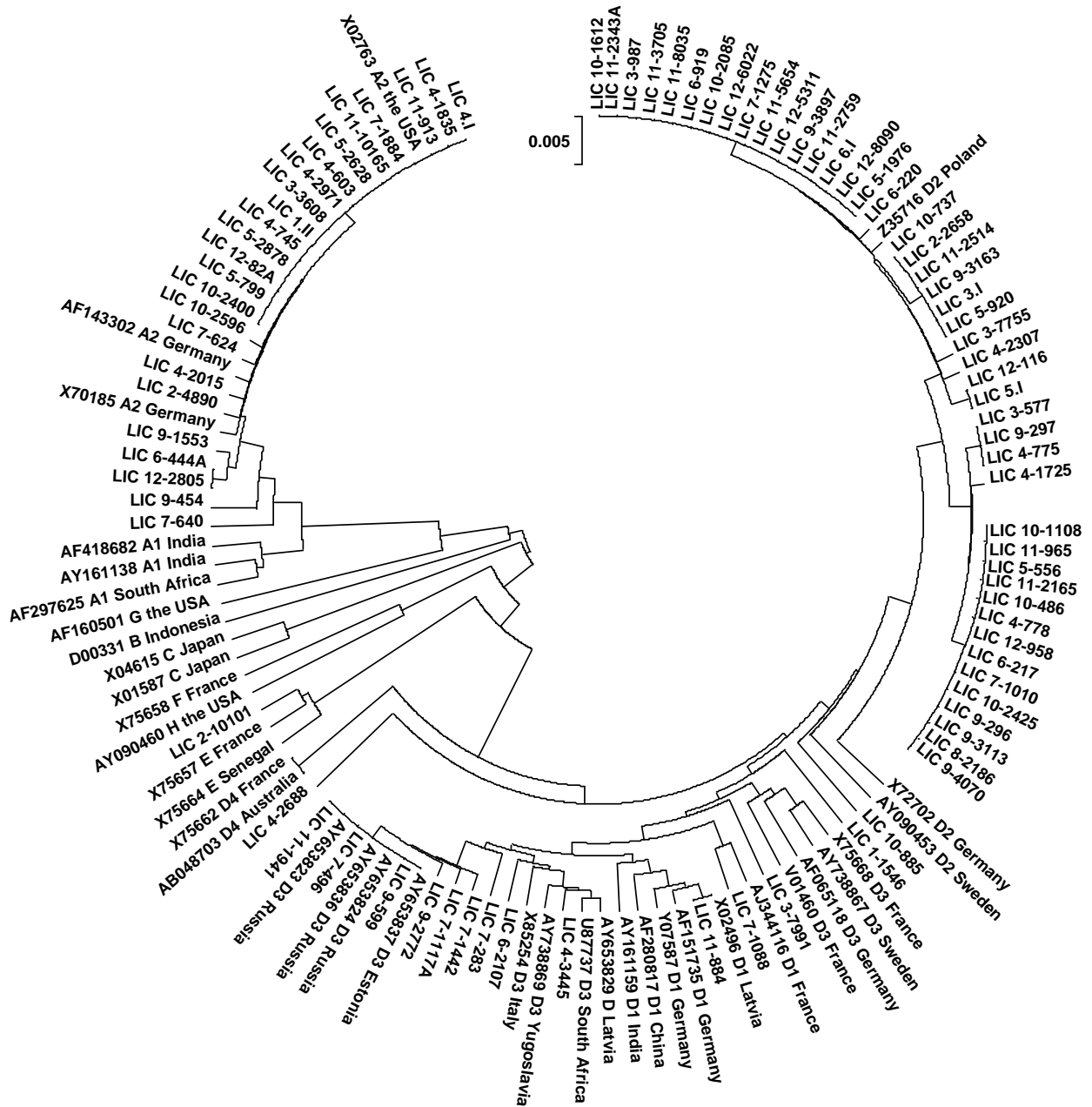


Fig. (1). The UPGMA phylogenetic tree based on sequences of fragments from HBV S-gene. Reference isolates are given specifying GeneBank access number, genotype and country of origin.

subgenotype. Nine isolates were similar to the subgenotype D3 genome.

Since monospecific subtype-determining antibodies were not available in Latvia, the HBV subtypes were determined by analysis of gene S sequences according to Magnius & Norder [7]. 26.32 % of our isolates belonged to the subtype *adw2*, and 10.53% and 47.37 % represented the subtypes *ayw2* and *ayw3*, respectively. In the case of one patient (1.32%), LIC1-1546, it was impossible to define subtype, since mutation affected aa 127 of HBsAg, which is subtype-

specific. In 22 cases (13.16%) of the genotype D and 7 cases (3.95%) of the genotype A, it was impossible to amplify the S region and determine subtypes, probably due to mutations within regions adjacent to primers (Table 4). All subgenotype A2 strains were of the *adw2* subtype. All subgenotype D2 strains belonged to the *ayw3* subtype. Two strains of the *ayw2* subtype were of the D1 subgenotype. Eight strains with the subtype *ayw2* and one of the *ayw3* subtype belonged to the D3 subgenotype. The distribution of genotypes and subtypes within different groups of patients is listed in Table 4.

Sequence Variation in the HBV Genomes

It is known that mutations appear in different regions of the HBV genome, and probably each stage of hepatitis B

infection may be affected by the appearance of specific mutations [15]. Thorough analysis of major HBV regions EnhII/BCP/X-preC/C and preS/S of three groups of patients

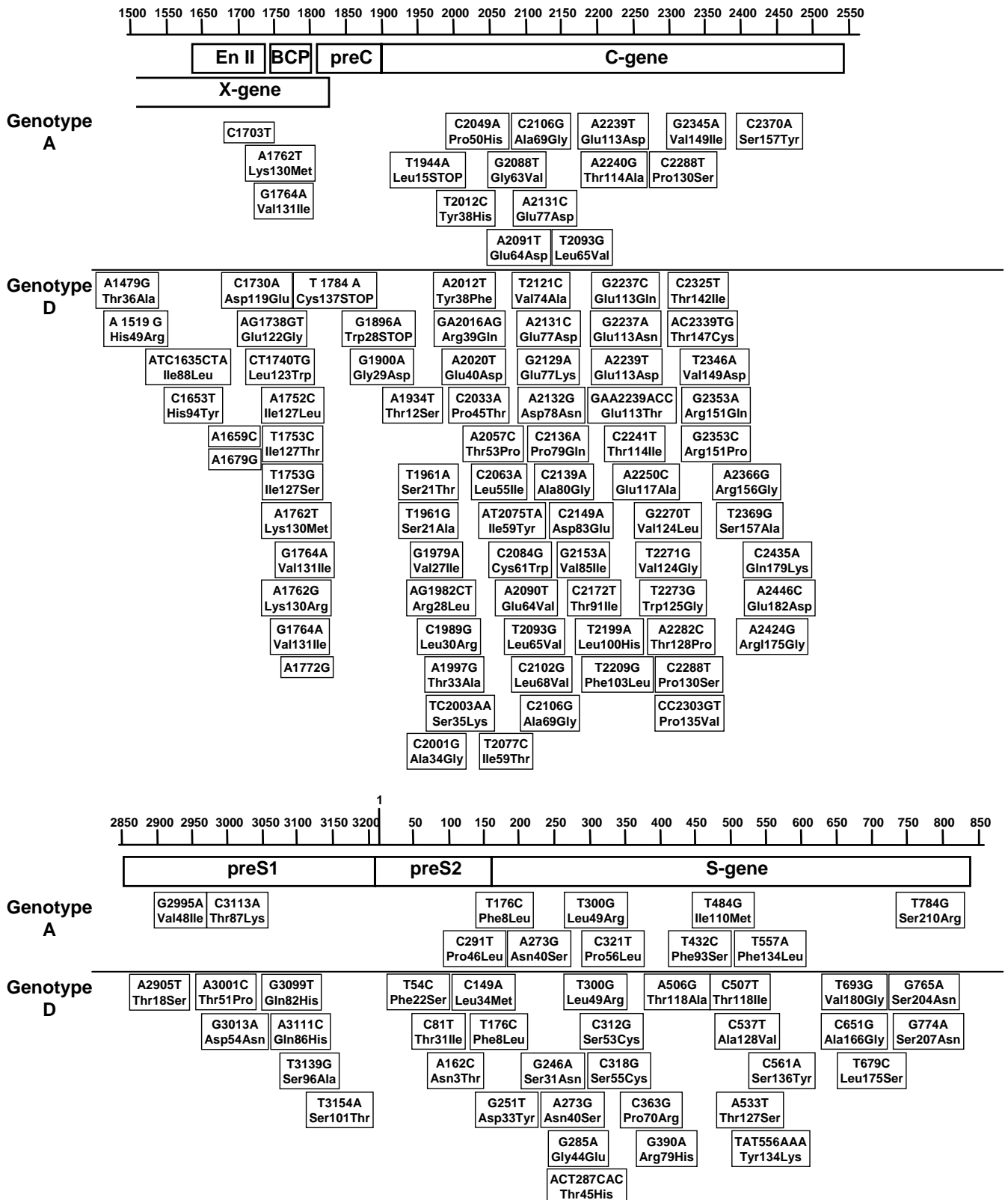


Fig. (2). Mutations discovered in tested HBV isolates.

was performed with the intention of revealing an association of the appearing mutations with the clinical course of HBV infection. Sequences of analyzed HBV regions compared with reference HBV sequences are given in Appendix. Detailed analysis of sequenced HBV genes is presented in Figs. (2 and 3). After genotype- and subtype-specific mutations, it reveals also the presence of silent and missense mutations.

Examining the distribution of mutations within the groups of patients and HBV genotypes (Fig. 3), we may conclude that, in general, in groups of CH patients and patients with other liver damage, mutations occur more frequently than in groups of AH patients. On the other hand, strains of genotype D contain more mutations when compared to strains of genotype A.

The distribution of the mutations on the genome of HBV correlates well enough with arrangement of the functionally and immunologically important regions. On the diagram we can see some hot spots corresponding to BCP, preC, MIR region in C-gene, and domain *a* in S-gene.

Gene X/EnII/BCP

The BCP mutations A1762T and G1764A, most frequently observed in HBV-infected patients with CH, HCC,

and FH, were found in one CH and six AH cases. Additional mutations at C1653T were found in two AH cases and T1753C in four AH and one CH case. These mutations may also be associated with the progression of liver disease to cirrhosis [16]. One AH isolate contained mutation at position T1753C in conjunction with the G1764A mutation. One CH isolate contained G1764T and C1765G mutations, which resulted in Val131Leu substitution in the gene X, in conjunction with the T1752C and T1753G mutations, which led to Ile127Thr substitution. The C1653T mutation led to the His94Tyr substitution in the immunodominant epitope of the protein X [17]. The A1752C mutation in patients with CH was reported also in conjunction with the 1762 and 1764 mutations and appeared to be associated with liver damage [18]. We found this mutation in five cases, and in all cases it was independent of other mutations. Two AH samples contained aa substitution Ile88Leu within the protein X. One AH isolate possessed T1784A mutation, which led to conversion of Cys137 to stop codon. Other individual AH isolates demonstrated A1479G mutation leading to the Thr36Ala substitution, A1519G mutation leading to His49Arg substitution, and C1730A mutations leading to Asp119Glu substitution. One AH isolate carried four mutations: AG1738GT and CT1740TG, which changed two aa residues: Glu122Gly and Leu123Trp.

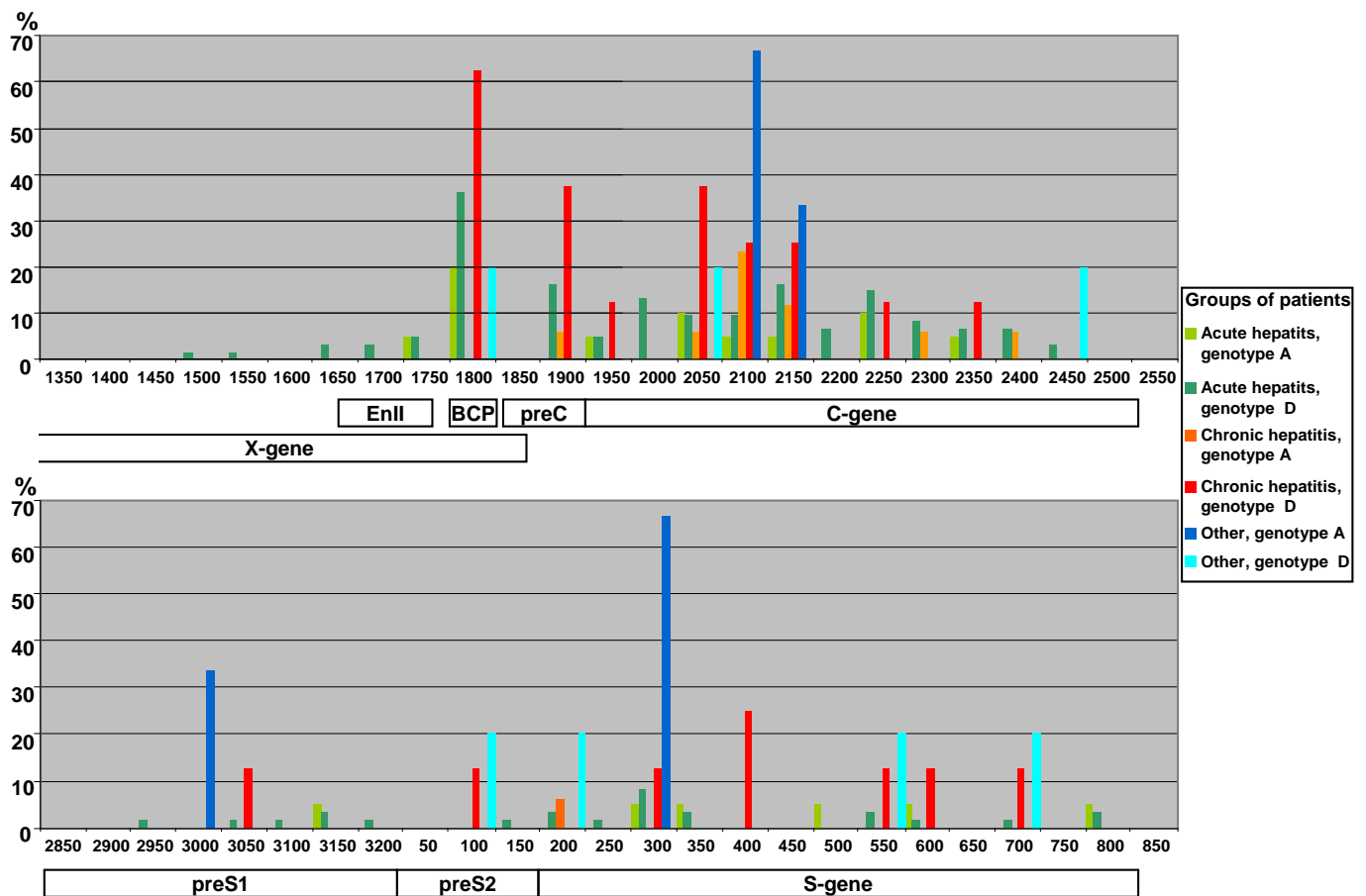


Fig. (3). Frequency of mutations discovered in HBV genome within groups of LIC patients and HBV genotypes.

It is interesting that the hepatic coma isolate possessed only few mutations, but two of them were located within the EnhII region.

PreC/C Gene

The preC stop codon mutation G1896A leading to substitution of the preC Trp -28 to UGA stop codon was detected in eight AH and three CH isolates. Absence of HBeAg in all except one patients' sera confirmed the observation that such mutations abort HBeAg expression and cause CH disease [19, 20]. The one CH and two AH isolates contained mutation at position 1900 (G→A), which changed aa 29 of the preC (Gly29Asp).

We detected numerous mutations within the HBc region 84–101, which overlaps the CTL epitopes (Fig. 2). Such mutations accumulate usually during periods with frequent hepatitis exacerbation [21,22]. Numerous unique mutations were found in the region 48–60 (Fig. 2), which were described as typical for FH [23–25] and severe exacerbation hepatitis B patients [26,27]. We detected also mutations in previously described mutation clustering regions at aa 21–34 [28] and 31–49 [29].

In two AH isolates we found substitutions within the HBc CTL epitope 18–27: Ser21Ala and Ser21Thr, which were not described before. HBc CTL epitope 18–27 mutations Ser21Asn/Ala/Val were found earlier in HLA-A2-positive CH patients [30]. In other HBc CTL epitopes: 63–71 and 141–149 [31], the following substitutions were found Gly63Val (twice in AH and CH patients), Glu64Asp (twice in LC and CH patients), and Glu64Val (once in AH patient). The Glu64Asp substitution was combined with two other mutations within the same epitope: Leu65Val and Ala69Gly. Mutations were found also within the C-terminal CTL epitope 141–149 (Fig. 2).

We found Pro130Ser substitution in two isolates: one AH and one CH. The aa 130 belongs to both T helper and B cell epitopes, and the Pro130Ile/Thr/Ser substitutions could be associated with exacerbation of chronic hepatitis [32,33] and HCC [34]. Such substitutions were detected in patients with or without HBeAg [35]. In our case the AH patient had no HBeAg in his sera, whereas CH patient had it. The Arg151Gln substitution, which could be regarded as common, was found once as well as Arg151Pro substitution, both in AH patients. We found a set of mutations within the T helper epitope 50–69 and within the B cell epitope 74–83, previously detected in patients seroconverted from HBe to anti-HBe [36]: Thr53Pro/Ala was found in two cases in CH and AH patients with anti-HBe in sera. The AH patient sample contained also Cys61Trp substitution. The Glu77Lys/Asp substitution was found in four AH patients, one of them was anti-HBe positive, another one was however HBeAg positive. The latter patient possessed also Asp78Asn substitution. The Pro79Gln mutation was found in one AH and two CH isolates, but the Ala80Gly substitution was found in an AH isolate, two of them were anti-HBe positive. A large number of mutations was found within the α 2a, α 2b, and α 5 helices of the HBc protein. The most frequent among them were: Tyr38His (two AH cases) and Tyr38Phe (one AH case), Glu40Asp (three AH and three CH cases), Glu113Gln (one AH and one CH case), Glu113Asp (four AH cases).

The Glu113Asn substitution was found once in AH isolate, Glu113Thr also once in AH isolate. The Thr114Ala/Ile and ValGlu124Leu substitutions were found in two AH cases.

In contrast to earlier observations [37], in all our studies we did not find any HBc deletion variants.

PreS/S Gene

Two mutations were found in the HBV preS1 region five times: G2995A that leads to Val48Ile substitution and G3001C - to Asp51His substitution. These mutations appeared in *adw2* patients, three times in AH and twice in CH patients. Single G2995A mutation was found in an LC isolate. One CH isolate possessed A3001C mutation leading to the Thr51Pro substitution. Besides this, the following individual mutations were observed: A2905T (Thr18Ser), G3013A (Asp54Asn), G3099T (Gln82His), A3111C (Gln86His), C3113A (Thr87Lys), T3139G (Ser96Ala), T3154A (Ser101Thr), all in AH isolates.

Only three mutations were found within the preS2 region: T54C (Phe22Ser) in a CH isolate, C81T (Thr31Ile) in a coma patient and C149A (Leu34Met) in an AH patient.

It was documented that chronic HBV carriers and HCC patients frequently have mutations encompassing residues 29–53 [38] of the S region. We found some mutations within this region, but only Leu49Arg substitution was found twice, but in AH isolates. Within the site of hypervariability located at position 134 [39] we found the following substitutions: Tyr134Phe/Asn/Trp/Ser or Phe134Tyr depending of the HBsAg subtype. At this position we found some novel substitutions - Tyr134Lys in a CH patient and Phe134Leu in an AH patient. Novel substitutions were also observed at positions Ile110Met, Ser136Tyr, and Thr118Ile, all in AH patients. Only Thr118Ala substitution was reported earlier [40]. We found the Leu175Ser substitution in the hepatic coma patient, although this mutation appeared before only in combination with a Phe179Tyr substitution [41].

HBV GENOTYPES IN LATVIA AND NEIGHBORING COUNTRIES

The present study is the first epidemiological attempt examining the prevalence of the HBV genotypes A and D and their mutational status in the wide group of patients admitted to the Infectology Center of Latvia. Our findings reflect the overall genotype and subtype distribution in the Latvian population: 71.52 % of genotype D and 27.81 % of genotype A, 0.87% of genotype E; 23.18 % of subtype *adw2*, 15.23 % of subtype *ayw2*, 39.07 % of subtype *ayw3*, 0.66% of subtype *ayw4* and 1.99 *adw3*. In one case, it was impossible to establish subtype due to unique Thr127Ser substitution at the subtype-specific residue, which has not been described previously. In our neighbor country, Poland, the genotype A is prevalent (74.1%) [42], but the genotype D is on the second place (20.7%). Almost all HBV genotype A cases are of the *adw2* subtype. As to the genotype D similar to major subtypes found in Latvia, they are represented by *ayw3* and *ayw2* variants. In Estonia, another neighboring country [43], the situation is similar to Latvia, with predominance of the genotype D and prevalence of the

subtypes *ayw3* and *adw2*. In contrast to Latvia, both neighboring countries show only a few cases of the *ayw4* subtype. Data from another neighbor, Belarus, show also prevalence of the D genotype (87.5%) over the A genotype (12.5%) [44]. The most widespread within the D genotype in Belarus is the D2 subgenotype (56.3%). In contrast to our data, the subgenotype D4 (3.1%) was also found. Our neighbor to the East, the Russian Federation, is very large and proportion of HBV genotypes varies strongly depending on geographic location [45]. Thus, the genotype D comprises 85 % among children in Moscow [45,46]. In Western Siberia, the genotype D constitutes even more than 96 % of the total. In contrast to Latvia, the *ayw2* subtype was clearly predominant over the *ayw3* subtype [47]. Our data also support the postulate, that genotype D is predominant in the Mediterranean region, Southern and Eastern Europe, in such countries as Hungary [48], Italy: 73% of the genotype D, and 26% of the genotype A [49,50], and Spain: 63,2% and 23.1%, respectively [51,52]. In most Western European countries, such as Belgium 54% of genotype A, and 31% of genotype D [53], the Netherlands 64%, and 21% respectively [54] genotype A dominates.

CONCLUSIONS

- This study indicates that the genotype D is a predominant genotype among patients in Latvia, as in most of

the neighboring countries The genotype D isolates are more heterogeneous than those of the genotype A.

- PreC/C mutants are prevalent among patients of Latvia. A never previously reported preC substitution Ser21Thr was found. At the same time we found no evidence for the existence of preC deletion variants.
- Comparison of the HBV sequences from the preS/S region with the strain of the same genotype revealed no deletions and relatively low number of mutations. At the same time we discovered S protein substitutions Tyr134Lys and Phe134Leu not previously described. Novel substitutions were observed also at positions Ile110Met, Thr118Ile and Ser136Tyr.

ACKNOWLEDGEMENTS

We thank Prof. Eva Stankevica and her group for oligonucleotide synthesis and automatic sequencing of HBV genome fragments, and Mrs. Viktorija Emeljanova for technical assistance.

This work was supported by National Program 10.0029.2 of the Latvian Council of Sciences, ERAF project 2DP/2.1.1.1.0/10/APIA/VIAA/164 and ESF project 1DP/1.1.1.2.0/09/APIA/VIAA/150.

Appendix

Aligned nucleotide and protein sequences of different regions of Hepatitis B genome obtained during the study. Sequences are ranged by genotype and diagnosis of patients: AH – acute hepatitis, CH – chronic hepatitis, COMPL – other complications. As a reference sequences are given for genotype A - X02763, for genotype D - V01460 and X02496, for genotype E - X75657. Synonymous nucleotides are shown as dots.

Appendix Part A. Sequences of HBV region preC/C obtained during the study.

	1820	1830	1840	1850	1860	1870	1880	1890	1900	1910	1920	1930
X02763	ATGCAACTTTTTCACCTCTGCCTAATCATCTCTGTACATGTCCCACTGTTCAAGCCTCCAAGCTGTGCCTTGGGTGGCTTTGGGGCATGGACATTGACCCCTATAAAGAATTGGAGCT											
AH 2-4890	M Q L F H L C L I I S C T C P T V Q A S K L C L G W L W G M D I D P Y K E F G A											
AH 1. I	M Q L F H L C L I I S C T C P T V Q A S K L C L G W L W G M D I D P Y K E F G A											
AH 4-1835	-----											
AH 4-2015G.....											
AH 5-2878	M Q L F H L C L I I S C T C P T V Q A S K L C L G W L W G M D I D P Y K E F G A											
AH 7-1884	M Q L F H L C L I I S C T C P T V Q A S K L C L G W L W G M D I D P Y K E F G A											
AH 7-624	-----											
AH 7-640T.....G.....											
AH 8-1993	M Q L F H L C L I I S C T C P T V Q A S K L C L G W L W G M D I D P Y K E F G A											
AH 10-2400	M Q L F H L C L I I S C T C P T V Q A S K L C L G W L W G M D I D P Y K E F G A											
AH 10-2596	M Q L F H L C L I I S C T C P T V Q A S K L C L G W L W G M D I D P Y K E F G A											
AH 11-2641	-----											
AH 12-21	-----											
AH 12-431	-----											

CH 1-6126	-----
CH 10-1982 M Q L F H L C L I I S C T C P T V Q A S K L C L G W L W G M D I D P Y K E F G A
CH 12-2805 M Q L F H L C L I I S C T C P T V Q A S K L C L G W L W G M D I D P Y K E F G A
CH 12-82	-----
CH 5-2628 M Q L F H L C L I I S C T C P T V Q A S K L C L G W L W G M D I D P Y K E F G A
CH 6-444A	-----
COMPL 4-2971 M Q L F H L C L I I S C T C P T V Q A S K L C L G W L W G M D I D P Y K E F G A
V01460	ATGCAACTTTTTCACCTCTGCCTAATCATCTCTTGTTCATGTCTACTGTTCAAGCCTCCAAGCTGTGCCTTGGGTGGCTTTGGGGCATGGACATCGACCCCTTATAAAGAATTTGGAGCT
X02496 M Q L F H L C L I I S C S C P T V Q A S K L C L G W L W G M D I D P Y K E F G AT.T.
AH 10-1108 M Q L F H L C L I I S C S C P T V Q A S K L C L G W L W G M D I D P Y K E F G AA.T.
AH 10-1612 M Q L F H L C L I I S C S C P T V Q A S K L C L G W L W G M D I D P Y K E F G AA.T.G.
CH 10-3329 M Q L F H L C L I I S C S C P T V Q A S K L C L G W L * G M D I D P Y K E F G AA.T.
AH 10-2085 M Q L F H L C L I I S C S C P T V Q A S K L C L G W L * G M D I D P Y K E F G AA.T.G.
AH 10-2425	-----
AH 10-737 M Q L F H L C L I I S C S C P T V Q A S K L C L G W L W G M D I D P Y K E F G AA.T.
AH 10-486 M Q L F H L C L I I S C S C P T V Q A S K L C L G W L W G M D I D P Y K E F G AA.T.
AH 11-1941 M Q L F H L C L I I S C S C P T V Q A S K L C L G W L * G M D I D P Y K E F G AA.T.
AH 11-2336A	-----
AH 11-2514	-----
AH 11-2524	-----
AH 11-2759 M Q L F H L C L I I S C S C P T V Q A S K L C L G W L W G M D I D P Y K E F G AA.T.G.
AH 11-2866	-----
AH 11-5664	-----
AH 11-965	-----
AH 12-438 M Q L F H L C L I I S C S C P T V Q A S K L C L G W L W G M D I D P Y K E F G AA.T.G.
AH 3-577	-----
AH 2-2658 M Q L F H L C L I I S C S C P T V Q A S K L C L G W L W G M D I D P Y K E F G AA.
AH 3-7991 M Q L F H L C L I I S C S C P T V Q A S K L C L G W L * D M D I D P Y K E F G AA.A.T.
AH 11-8035 M Q L F H L C L I I S C S C P T V Q A S K L C L G W L * G M D I D P Y K E F G AA.T.
AH 3-7755 M Q L F H L C L I I S C S C P T V Q A S K L C L G W L * G M D I D P Y K E F G AA.A.G.
AH 5-920 M Q L F H L C L I I S C S C P T V Q A S K L C L G W L W G M D I D P Y K E F G AA.
AH 3-987 M Q L F H L C L I I S C S C P T V Q A S K L C L G W L W G M D I D P Y K E F G AA.T.G.
AH 4-2307	-----
AH 4-3445 M Q L F H L C L I I S C S C P T V Q A S K L C L G W L W G M D I D P Y K E F G AA.T.
AH 4-775	-----

	1820	1830	1840	1850	1860	1870	1880	1890	1900	1910	1920	1930
V01460	ATGCAACTTTTTCACCTCTGCCTAATCATCTCTTGTTCATGCTCTACTGTTCAAGCCTCCAAGCTGTGCCTTGGGTGGCTTTGGGGCATGGACATCGACCCCTTATAAAGAATTGGAGCT											
X02496	M Q L F H L C L I I S C S C P T V Q A S K L C L G W L W G M D I D P Y K E F G A											
AH 4-778	-----											
AH 6.I	-----											
AH 5-556A.....T.....											
AH 6-1289	M Q L F H L C L I I S C S C P T V Q A S K L C L G W L W G M D I D P Y K E F G A											
AH 6-217	-----											
AH 6-220A.....T.....G.....											
AH 6-2394	M Q L F H L C L I I S C S C P T V Q A S K L C L G W L W G M D I D P Y K E F G A											
AH 6-2616A.....T.....											
AH 7-1010	M Q L F H L C L I I S C S C P T V Q A S K L C L G W L * G M D I D P Y K E F G A											
AH 7-1275A.....T.....											
AH 7-1440	-----											
AH 7-1442	-----											
AH 7-1501A.....A.....T.....											
AH 7-283	M Q L F H L C L I I S C S C P T V Q A S K L C L G W L * G M D I D P Y K E F G A											
AH 7-3552A.....A.....T.....											
AH 8-1853	M Q L F H L C L I I S C S C P T V Q A S K L C L G W L * G M D I D P Y K E F G A											
AH 8-2186A.....T.....											
AH 9-2772A.A.....T.....											
AH 9-296	M Q L F H L C L I I S C S C P T V Q A S K L C L G W L * D M D I D P Y K E F G A											
AH 9-3113A.....T.....											
AH 9-3163A.....T.....											
AH 9-4070A.....T.....											
CH 1-1546	M Q L F H L C L I I S C S C P T V Q A S K L C L G W L * G M D I D P Y K E F G A											
CH 1-4936A.A.....T.....											
CH 10-3335A.....T.....											
CH 11-2343A	M Q L F H L C L I I S C S C P T V Q A S K L C L G W L * G M D I D P Y K E F G A											
CH 11-2866	-----											
CH 11-5664	-----											
CH 11-884	-----											
CH 12-2178	-----											
CH 12-958	-----											
CH 6-2625	-----											
CH 7-1117AT.....											
CH 7-2508	M Q L F H L C L I I S C S C P T V Q A S K L C L G W L W G M D I D P Y K E F G A											
CH 9-599T.....											
CH 9-7181	M Q L F H L C L I I S C S C P T V Q A S K L C L G W L W G M D I D P Y K E F G A											
COMPL 11-3028A.....T.....											
COMPL 11-3179A.....T.....											
COMPL 4-1725A.....T.....											
COMPL 7-2083	M Q L F H L C L I I S C S C P T V Q A S K L C L G W L W G M D I D P Y K E F G A											
X75657	ATGCAACTTTTTCACCTCTGCCTAATCATCTCTTGTTCATGCTCTACTGTTCAAGCCTCCAAGCTGTGCCTTGGGTGGCTTTGGGGCATGGACATCGACCCCTTATAAAGAATTGGAGCT											
AH 2-10101	M Q L F H L C L I I S C S C P T V Q A S K L C L G W L W G M D I D P Y K E F G A											

	1940	1950	1960	1970	1980	1990	2000	2010	2020	2030	2040	2050
X02763	ACTGTGGAGTTACTCTCGTTTTTGCCTTCTGACTTCTTCTCCGTCAGAGATCTCTAGACACCGCCTCAGCTCTGTATCGAGAAGCCTTAGAGTCTCCTGAGCATTGCTCACCTCAC											
AH 2-4890	T V E L L S F L P S D F F P S V R D L L D T A S A L Y R E A L E S P E H C S P H											
AH 1. I	T V E * L S F L P S D F F P S V R D L L D T A S A L Y R E A L E S P E H C S P H											
AH 4-1835	-----											
AH 4-2015	-----											
AH 5-2878	T V E L L S F L P S D F F P S V R D L L D T A S A L Y R E A L E S P E H C S P H											
AH 7-1884	T V E L L S F L P S D F F P S V R D L L D T A S A L Y R E A L E S P E H C S P H											
AH 7-624	-----											
AH 7-640	-----											
AH 8-1993	T V E L L S F L P S D F F P S V R D L L D T A S A L Y R E A L E S P E H C S P H											
AH 10-2400	-----											
AH 10-2596	T V E L L S F L P S D F F P S V R D L L D T A S A L H R E A L E S P E H C S P H											
AH 11-2641	-----											
AH 12-21	-----											
AH 12-431	-----											
CH 1-6126	-----											
CH 10-1982	-----											
CH 12-2805	T V E L L S F L P S D F F P S V R D L L D T A S A L Y R E A L E S P E H C S P H											
CH 12-82	-----											
CH 5-2628	-----											
CH 6-444A	-----											
COMPL 4-2971	-----											
V01460	ACTGTGGAGTTACTCTCGTTTTTGCCTTCTGACTTCTTCTCCGTCAGAGATCTCTAGATACCGCCTCAGCTCTGTATCGGGAAGCCTTAGAGTCTCCTGAGCATTGCTCACCTCAC											
X02496	T V E L L S F L P S D F F P S V R D L L D N A S A L Y R E A L E S P E H C S P H											
AH 10-1108	-----											
AH 10-1612	-----											
CH 10-3329	-----											
AH 10-2085	-----											
AH 10-2425	-----											
AH 10-737	-----											
AH 10-486	-----											
AH 11-1941	T.C. S V E L L S F L P S D F F P S V R D L L D T A S A L Y R D A L E S P E H C S P H											
AH 11-2336A	-----											
AH 11-2514	-----											
AH 11-2524	-----											
AH 11-2759	-----											
AH 11-2866	-----											
AH 11-5664	-----											
AH 11-965	-----											
AH 12-438	-----											
AH 3-577	-----											
AH 2-2658	-----											
AH 3-7991	T. S V E L L S F L P A D F F P S I R D L L D T A K A L F Q D A L E S P E H C S P H											
AH 11-8035	-----											
AH 3-7755	-----											
AH 5-920	-----											
AH 3-987	-----											
AH 4-2307	-----											
AH 4-3445	-----											
AH 4-775	-----											

	1940	1950	1960	1970	1980	1990	2000	2010	2020	2030	2040	2050
V01460	ACTGTGGAGTTACTCTCGTTTTGCGCTTCTGACTTCTTTCCTTCAGTACGAGATCTTCTAGATACCGCCTCAGCTCTGTATCGGGAAGCCTTAGAGTCTCCTGAGCATTGTTCACTCAC											
X02496	T V E L L S F L P S D F F P S V R D L L D T A S A L Y R E A L E S P E H C S P H											
AH 4-778	T V E L L S F L P S D F F P S V R D L L D N A S A L Y R E A L E S P E H C S P H											
AH 6.I	-----											
AH 5-556	-----											
AH 6-1289	T V E L L S F L P S D F F P S V R D L L D T A S A L Y R E A L E S P E H C S P H											
AH 6-217	-----											
AH 6-220	-----											
AH 6-2394	T V E L L S F L P S D F F P S V R D L L D T A S A L Y R E A L E S T E H C S P H											
AH 6-2616	T V E L L S F L P S D F F P S V R D L L D T A S A L Y R E A L E S P E H C S P H											
AH 7-1010	T V E L L S F L P S D F F P S V R D L L D T A S A L Y R D A L E S P E H C S P H											
AH 7-1275	T V E L L S F L P S D F F P S V R D L L D T A S A L Y R E A L E S P E H C S P H											
AH 7-1440	-----											
AH 7-1442	-----											
AH 7-1501	-----											
AH 7-283	T V E L L S F L P S D F F P S V R D L L D T A S A L Y R E A L E S P E H C S P H											
AH 7-3552	-----											
AH 8-1853	T V E L L S F L P S D F F P S V R D L L D T A S A L Y R E A L E S P E H C S P H											
AH 8-2186	T V E L L S F L P S D F F P S V R D L L D T A S A L Y R E A L E S P E H C S P H											
AH 9-2772	T A A . C A T A C A											
AH 9-296	S V E L L S F L P T D F F P S V R D L L D T A T A L Y R D A L E S P E H C T P H											
AH 9-3113	T V E L L S F L P S D F F P S V R D L L D T A S A L Y R E A L E S P E H C S P H											
AH 9-3163	T V E L L S F L P S D F F P S V R D L L D T A S A L Y R E A L E S P E H C S P H											
AH 9-4070	T V E L L S F L P S D F F P S V R D L L D T A S A L Y R E A L E S P E H C S P H											
CH 1-1546	T . C A T A											
CH 1-4936	S V E L L S F L P S D F Y P S V R D L L D T A S A L Y R D A L E S P E H C S P H											
CH 10-3335	. . C T A											
CH 11-2343A	T V E L L S F L P S D F F P S V R D L L D T A S A L Y R D A L E S P E H C S P H											
CH 11-2866	-----											
CH 11-5664	-----											
CH 11-884	-----											
CH 12-2178	-----											
CH 12-958	-----											
CH 6-2625	-----											
CH 7-1117A	. . C T A											
CH 7-2508	T V E L L S F L P S D F F P S V R D L L D T A S A L Y R D A L E S P E H C S P H											
CH 9-599	. . C T A											
CH 9-7181	T V E L L S F L P S D F F P S V R D L L D T A S A L Y R D A L E S P E H C S P H											
COMPL 11-3028	-----											
COMPL 11-3179	T V E L L S F L P S D F F P S V R D L L D T G S A L Y R E A L E S P E H C S P H											
COMPL 4-1725	T V E L L S F L P S D F F P S V R D L L D T A S A L Y R E A L E S P E H C S P H											
COMPL 7-2083	T V E L L S F L P S D F F P S V R D L L D T A S A L Y R E A L E S P E H C S P H											
X75657	ACTGTGGAGTTACTCTCGTTTTGCGCTTCTGACTTCTTTCCTTCAGTAAAGAGATCTTCTAGATACCGCCTCAGCTCTGTATCGGGATGCCTTAGAGTCTCCTGAGCATTGTTCACTCAC											
AH 2-10101	T V E L L S F L P S D F F P S V L D L L D T A S A L Y R D A L E S P E H C S P H											

	2060	2070	2080	2090	2100	2110	2120	2130	2140	2150	2160	2170
X02763	CATACTGCACCTCAGGCAAGCAATCTCTGCTGGGGGGAATGATGACTCTAGCTACCTGGGGGGTAAATAATTGGAAGATCCAGCATCTAGGAGCTCTTAGTAAATATGTTAATACT											
AH 2-4890	H T A L R Q A I L C W G E L M T L A T W V G N N L E D P A S R D L V V N Y V N T						C		C	A	C	
AH 1.1	H T A L R Q A I L C W G E L M T L A T W V G N N L E D P A S R D L V V N Y V N T			T					C	A	C	
AH 4-1835	H T A L R Q A I L C W V E L M T L A T W V G N N L E D P A S R D L V V N Y V N T									A	C	
AH 4-2015	H T A L R Q A I L C W G E L M T L A T W V G N N L E D P A S R D L V V N Y V N T								C	A		
AH 5-2878	H T A L R Q A I L C W G E L M T L A T W V G N N L E D P A S R D L V V N Y V N T											
AH 7-1884	H T A L R Q A I L C W G E L M T L A T W V G N N L E D P A S R D L V V N Y V N T								C	A	G	C
AH 7-624	H T A L R Q A I L C W G E L M T L A T W V G N N L E D P A S R D L V V N Y V N T								C	A	C	
AH 7-640	H T A L R Q A I L C W G E L M T L A T W V G N N L E D P A S R D L V V N Y V N T								C	A	C	
AH 8-1993	H T A L R Q A I L C W G E L M T L A T W V G N N L D D P A S R D L V V N Y V N T								C	A	G	C
AH 10-2400	H T A L R Q A I L C W G E L M T L A T W V G N N L E D P A S R D L V V N Y V N T								C	A	C	
AH 10-2596	H T A L R Q A I L C W G E L M T L A T W V G N N L E D P A S R D L V V N Y V N T								C	C	A	C
AH 11-2641	H T A L R Q A I L C W G E L M T L A T W V G N N L E D P A S R D L V V N Y V N T								C	A	C	
AH 12-21	H T A L R Q A I L C W G E L M T L A T W V G N N L E D P A S R D L V V N Y V N T								C	A	C	
AH 12-431	H T A L R Q A I L C W G E L M T L A T W V G N N L E D P A S R D L V V N Y V N T								C	A	C	
CH 1-6126	H T A L R Q A I L C W G E L M T L A T W V G N N L E D P A S R D L V V N Y V N T			TG	G				C	A	C	
CH 10-1982	H T A L R Q A I L C W G D V M T L G T W V G N N L D D P A S R D L V V N Y V N T			TG	G				C	A	C	
CH 12-2805	H T A L R Q A I L C W G D V M T L G T W V G N N L E D P A S R D L V V N Y V N T								C	A	C	
CH 12-82	H T A L R Q A I L C W V E L M T L A T W V G N N L E D P A S R D L V V N Y V N T								C	A	C	
CH 5-2628	H T A L R Q A I L C W G E L M T L A T W V G N N L E D P A S R D L V V N Y V N T			TG	G				C	A	C	
CH 6-444A	H T A L R Q A I L C W G D V M T L G T W V G N N L E D P A S R D L V V N Y V N T								C	A	C	
CH 6-444A	H T A L R Q A I L C W G E L M T L A T W V G N N L E D P A S R D L V V N Y V N T			T					G	C	A	C
COMPL 4-2971	H T A L R Q A I L C W G E L M T L A T W V G N N L E D P A S R D L V V N Y V N T			TG	G				C	A	C	
V01460	CATACTGCACCTCAGGCAAGCAATCTCTGCTGGGGGGAATGATGACTCTAGCTACCTGGGGGGTAAATAATTGGAAGATCCAGCATCTAGGAGCTCTTAGTAAATATGTTAATACT											
X02496	H T A L R Q A I L C W G E L M T L A T W V G V N L E D P A S R D L V V S Y V N T								ATA	C	G	
AH 10-1108	H T A L R Q A I L C W G E L M T L A T W V G G N L E D P I S R D L V V S Y V N T			A	G					G		
AH 10-1612	H T A L R Q A I L C W G E L M T L A T W V G V N L E D P A S R D L V V S Y V N T								A	G		
CH 10-3329	H T A L R Q A I L C W G E L M T L A T W V G V N L E D P A S R D L V V S Y V N T								C	A	G	
AH 10-2085	H T A L R Q A I L C W G E L M T L A T W V G V N L E D P A S R D L V V S Y V N T								A	G		
AH 10-2425	H T A L R Q A I L C W G E L M T L A T W V G V N L E D P A S R D L V V S Y V N T								A	G		
AH 10-737	H T A L R Q A I L C W G E L M T L A T W V G V N L E D P A S R D L V V S Y V N T								A	G		
AH 10-486	H T A L R Q A I L C W G E L M T L A T W V G V N L E D P A S R D L V V S Y V N T								A	G		
AH 11-1941	H T A L R Q A I L C W G E L M T L A T W V G V N L E D P A S R D L V V S Y V N T			T	G				A	G		
AH 11-2336A	H T A L R Q A I L C W G E L M T L A T W V G V N L E D Q A S R D L V V S Y V N T								A	G		
AH 11-2514	H T A L R Q A I L C W G E L M T V A T W V G V N L E D P A S R D L V V S Y V N T								A	G		
AH 11-2524	H T A L R Q A I L C W G E L M T L A T W V G V N L E D P A S R D L V V S Y V N T								A	G		
AH 11-2759	H T A L R Q A I L C W G E L M T L A T W V G V N L E D P A S R D L V V S Y V N T								A	G		
AH 11-2866	H T A L R Q A I L C W G E L M T L A T W V G V N L E D P A S R D L V V S Y V N T								A	A	G	
AH 11-5664	H T A L R Q A I L C W G E L M T L A T W V G V N L E D P A S R D L V V S Y V N T								A	G		
AH 11-965	H T A L R Q A I L C W G E L M T L A T W V G V N L E D P A S R D L V V S Y V N T								A	G		
AH 11-965	H P A L R Q A I L W W G E L M T L A T W V G V N L E D P G S R D L V V S Y V N T								GA	C	G	
AH 12-438	H T A L R Q A I L C W G E L M T L A T W V G V N L E D P A S R D L V V S Y V N T								A	G		
AH 3-577	H T A L R Q A I L C W G E L M T L A T W V G V N L E D P A S R D L V V S Y V N T								A	G		
AH 2-2658	H T A L R Q A I L C W G E L M T L A T W V G V N L E D P A S R D L V V S Y V N T								A	G		
AH 3-7991	H T A L R Q A I L C W G E L M T L A T W V G V N L E D P A S R D L V V S Y V N T			A	TA				C	T		
AH 3-7991	H T A I R Q A Y L C W G E L M T L A T W V G A N L D D P A S R E Q I V N Y V N I								A	A	A	
AH 11-8035	H T A L R Q A I L C W G E L M T L A T W V G V N L E D P A S R D L V V S Y V N T								A	G		
AH 3-7755	H T A L R Q A I L C W G E L M T L A T W V G V N L E D P A S R D L V V S Y V N T								A	G		
AH 5-920	H T A L R Q A I L C W G E L M T L A T W V G V N L E D P A S R D L V V S Y V N T								A	G		
AH 3-987	H T A L R Q A I L C W G E L M T L A T W V G V N L E D P A S R D L V V S Y V N T								A	G		
AH 4-2307	H T A L R Q A I L C W G E L M T L A T W V G G N L E D P I S R D L V V S Y V N T								A	G		
AH 4-3445	H T A L R Q A I L C W G E L M T L A T W V G V N L E D P A S R D L V V S Y V N T								A	G		
AH 4-775	H T A L R Q A I L C W G E L M T L A T W V G V N L E D P A S R D L V V S Y V N T								A	G		

	2180	2190	2200	2210	2220	2230	2240	2250	2260	2270	2280	2290
X02763	AACTGGGTTTAAAGATCAGGCAACTATTTGGTTTCATATCTTGCCCTACTTTTGGGAGAGAGACTGACTTGAATATTTGGTCTCTTTCCGAGTGTGGATTCCGACTCCTCCAGCC											
AH 2-4890	N V G L K I R Q L L W F H I S C L T F G R E T V L E Y L V S F G V W I R T P P A	A	G									
AH 1.1	N M G L K I R Q L L W F H I S C L T F G R E T V L E Y L V S F G V W I R T P P A	A	G									
AH 4-1835	N M G L K I R Q L L W F H I S C L T F G R E T V L E Y L V S F G V W I R T P P A	A	G									
AH 4-2015	N M G L K I R Q L L W F H I S C L T F G R E T V L E Y L V S F G V W I R T P P A	A	G									
AH 5-2878	N V G L K I R Q L L W F H I S C L T F G R E T V L E Y L V S F G V W I R T P P A											
AH 7-1884	N M G L K I R Q L L W F H I S C L T F G R E T V L E Y L V S F G V W I R T P P A	A										
AH 7-624	N M G L K I R Q L L W F H I S C L T F G R E T V L E Y L V S F G V W I R T P P A	A										
AH 7-640	N M G L K I R Q L L W F H I S C L T F G R E T V L E Y L V S F G V W I R T P P A	A	G				C					
AH 8-1993	N M G L K I R Q L L W F H I S C L T F G R E T V L E Y L V S F G V W I R T P P A	A										
AH 10-2400	N M G L K I R Q L L W F H I S C L T F G R E T V L E Y L V S F G V W I R T P P A	A						T				
AH 10-2596	N M G L K I R Q L L W F H I S C L T F G R E T V L E Y L V S F G V W I R T P P A	A	G									
AH 11-2641	N M G L K I R Q L L W F H I S C L T F G R E T V L E Y L V S F G V W I R T P P A	A	G									
AH 12-21	N M G L K I R Q L L W F H I S C L T F G R E T V L E Y L V S F G V W I R T P P A	A										
AH 12-431	N M G L K I R Q L L W F H I S C L T F G R E T V L E Y L V S F G V W I R T P P A	A	C	G								
CH 1-6126	N M G L K I R Q L L W F H I S C L T F G R E T V L E Y L V S F G V W I R T P P A	A	C				C					
CH 10-1982	N M G L K I R Q L L W F H I S C L T F G R E T V L E Y L V S F G V W I R T P P A	A	C							T		
CH 12-2805	N M G L K I R Q L L W F H I S C L T F G R E T V L E Y L V S F G V W I R T P S A	A	G									
CH 12-82	N M G L K I R Q L L W F H I S C L T F G R E T V L E Y L V S F G V W I R T P P A	A	G									
CH 5-2628	N M G L K I R Q L L W F H I S C L T F G R E T V L E Y L V S F G V W I R T P P A	A	C									
CH 6-444A	N M G L K I R Q L L W F H I S C L T F G R E T V L E Y L V S F G V W I R T P P A	A	G									
COMPL 4-2971	N M G L K I R Q L L W F H I S C L T F G R E T V L E Y L V S F G V W I R T P P A	A	C									
V01460	AATATGGGCTAAAGTTCAGGCAACTCTTTGGTTTCACATTTCTTGCTCACTTTTGGGAGAGAAACAGTTATAGAGTATTTGGTGTCTTTCCGAGTGTGGATTCCGACTCCTCCAGCC											
X02496	N M G L K F R Q L L W F H I S C L T F G R E T V I E Y L V S F G V W I R T P P A	A	A							T		
AH 10-1108	N M G L K F R Q L L W F H I S C L T F G R E T V I E Y L V S F G V W I R T P P A		T	A				G	C			
AH 10-1612	N M G L K F R Q L L W F H I S C L T F G R E T V I E Y L V S F G V W I R T P P A	T	A					G	C			
CH 10-3329	N M G L K F R Q L L W F H I S C L T F G R E T V I E Y L V S F G V W I R T P P A		A					A				
AH 10-2085	N M G L K F R Q L L W F H I S C L T F G R E T V I E Y L V S F G V W I R T P P A	T	A	T				G	C			
AH 10-2425	N M G L K F R Q L L W F H I S C L T F G R E T V I E Y L V S F G V W I R T P P A		T					G	C		G	G
AH 10-737	N M G L K F R Q L L W F H I S C L T F G R E T V I E Y L V S F G V W I R T P P A	T	A	T				G	C			C
AH 10-486	N M G L K F R Q L L W F H I S C L T F G R E T V I E Y L V S F G V W I R P P P A		T					G	C		T	
AH 11-1941	N M G L K F R Q L L W F H I S C L T F G R E T V I E Y L V S F G L W I R T P P A		A									
AH 11-2336A	N M G L K F R Q L L W F H I S C L T F G R Q T V I E Y L V S F G V W I R T P P A		A				C		A			C
AH 11-2514	N M G L K F R Q H L W F H I S C L T F G R E T V I E Y L V S F G V W I R T P P A	T	A	AT				G	C			
AH 11-2524	N M G L K F R Q L L W F H I S C L T F G R E T V I E Y L V S F G V W I R T P P A		T	A				G	C			
AH 11-2759	N M G L K F R Q L L W F H I S C L T F G R E T V I E Y L V S F G V W I R T P P A	T	A	T				G	C			
AH 11-2866	N M G L K F R Q L L W F H I S C L T F G R E T V I E Y L V S F G V W I R T P P A		T					G	C			
AH 11-5664	N M G L K F R Q L L W F H I S C L T F G R E T V I E Y L V S F G V W I R T P P A		T					G	C	T		
AH 11-965	N M G L K F R Q L L W F H I S C L T F G R E T V I E Y L V S F G V W I R T P P A		T				C	A	T	G	C	
AH 12-438	N M G L K F R Q L L W F H I S C L T F G R E T V I E Y L V S F G V W I R T P P A	T	A	T				G	C			
AH 3-577	N M G L K F R Q L L W F H I S C L T F G R E T V I E Y L V S F G V W I R T P P A		T					G	C			
AH 2-2658	N M G L K F R Q L L W F H I S C L T F G R E T V I E Y L V S F G V W I R T P P A	T	A	T				G	C			
AH 3-7991	N M G L K F R Q L L W F H I S C L T F G R T V I E Y L V S F G V W I R T P P A		A				ACC		A			
AH 11-8035	N M G L K F R Q L L W F H I S C L T F G R E T V I E Y L V S F G V W I R T P P A		T					G	C			
AH 3-7755	N M G L K F R Q L L W F H I S C L T F G R E T V I E Y L V S F G V W I R T P P A	T	A	T				G	C			
AH 5-920	N M G L K F R Q L L W F H I S C L T F G R E T V I E Y L V S F G V W I R T P P A	T	A	T				G	C			
AH 3-987	N M G L K F R Q L L W F H I S C L T F G R E T V I E Y L V S F G V W I R T P P A		A					G	C			
AH 4-2307	N M G L K F R Q L L W F H I S C L T F G R E T V I E Y L V S F G V W I R T P P A	T	A					G	C			
AH 4-3445	N M G L K F R Q L L W F H I S C L T F G R E T V I E Y L V S F G V W I R T P P A		A							C		
AH 4-775	N M G L K F R Q L L W F H I S C L T F G R E T V I E Y L V S F G V W I R T P P A		T					G	C			

	2180	2190	2200	2210	2220	2230	2240	2250	2260	2270	2280	2290
V01460	AATATGGGCCCTAAAGTTCAGGCAACTCTGTGGTTTCACATTTCTGTCTCACTTTTGGGAGAGAAAACAGTTATAGAGTATTTGGTGTCTTTCCGAGTGTGGATTCCGACTCCTCCAGCT											
X02496	NMGLKFRQLLWPFHISCLTFFGRETVIEEYLVSFQGVWIRTPPA A A T											
AH 4-778	NMGLKFRQLLWPFHISCLTFFGRETVIEEYLVSFQGVWIRTPPA G.C.											
AH 6.I	NMGLKFRQLLWPFHISCLTFFGRETVIEEYLVSFQGVWIRTPPA T.A.T.G.C.											
AH 5-556	NMGLKFRQLLWPFHISCLTFFGRETVIEEYLVSFQGVWIRTPPA A.A.T.G.C.											
AH 6-1289	NMGLKFRQLLWPFHISCLTFFGRETVIEEYLVSFQGVWIRTPPA A.A.											
AH 6-217	NMGLKFRQLLWPFHISCLTFFGRETVIEEYLVSFQGVWIRTPPA T.A.T.G.C.											
AH 6-220	NMGLKFRQLLWPFHISCLTFFGRETVIEEYLVSFQGVWIRTPPA A.A.G.C.											
AH 6-2394	NMGLKFRQLLWPFHISCLTFFGRETVIEEYLVSFQGVWIRTPPA T.T.G.C.											
AH 6-2616	NMGLKFRQLLWPFHISCLTFFGRETVIEEYLVSFQGVWIRTPPA A.A.											
AH 7-1010	NMGLKFRQLLWPFHISCLTFFGRETVIEEYLVSFQGVWIRTPPA G.A.G.C.											
AH 7-1275	NMGLKFRQLLWPFHISCLTFFGRETVIEEYLVSFQGVWIRTPPA T.A.T.G.C.											
AH 7-1440	NMGLKFRQLLWPFHISCLTFFGRETVIEEYLVSFQGVWIRTPPA T.T.G.C.											
AH 7-1442	NMGLKFRQLLWPFHISCLTFFGRETVIEEYLVSFQGVWIRTPPA G.A.A.C.											
AH 7-1501	NMGLKFRQLLWPFHISCLTFFGRETVIEEYLVSFQGVWIRTPPA T.T.G.C.											
AH 7-283	NMGLKFRQLLWPFHISCLTFFGRETVIEEYLVSFQGVWIRTPPA A.A.											
AH 7-3552	NMGLKFRQLLWPFHISCLTFFGRETVIEEYLVSFQGVWIRTPPA T.T.G.C.											
AH 8-1853	NMGLKFRQLLWPFHISCLTFFGRETVIEEYLVSFQGVWIRTPPA T.A.T.G.C.											
AH 8-2186	NMGLKFRQLLWPFHISCLTFFGRETVIEEYLVSFQGVWIRTPPA T.A.T.G.C.											
AH 9-2772	NMGLKFRQLLWPFHISCLTFFGRETVIEEYLVSFQGVWIRTPPA A.A.G.T.TC.CT.A.											
AH 9-296	NMGLKFRQLLWPFHISCLTFFGRETVIEEYLVSFQGVWIRTPPA T.T.G.C.											
AH 9-3113	NMGLKFRQLLWPFHISCLTFFGRETVIEEYLVSFQGVWIRTPPA T.T.G.C.											
AH 9-3163	NMGLKFRQLLWPFHISCLTFFGRETVIEEYLVSFQGVWIRTPPA T.A.T.G.C.											
AH 9-4070	NMGLKFRQLLWPFHISCLTFFGRETVIEEYLVSFQGVWIRTPPA T.T.G.C.											
CH 1-1546	HMG LKFRQLLWPFHISCLTFFGRQT VIEEYLVSFQGVWIRTPPA C.T.A.C.A.T.											
CH 1-4936	NMGLKFRQLLWPFHISCLTFFGRETVIEEYLVSFQGVWIRTPPA A.A.											
CH 10-3335	NMGLKFRQLLWPFHISCLTFFGRETVIEEYLVSFQGVWIRTPPA A.A.											
CH 11-2343A	NMGLKFRQLLWPFHISCLTFFGRETVIEEYLVSFQGVWIRTPPA T.A.T.G.C.											
CH 11-2866	NMGLKFRQLLWPFHISCLTFFGRETVIEEYLVSFQGVWIRTPPA T.T.G.C.											
CH 11-5664	NMGLKFRQLLWPFHISCLTFFGRETVIEEYLVSFQGVWIRTPPA T.T.G.C.T.											
CH 11-884	NMGLKFRQLLWPFHISCLTFFGRETVIEEYLVSFQGVWIRTPPA A.A.T.											
CH 12-2178	NMGLKFRQLLWPFHISCLTFFGRETVIEEYLVSFQGVWIRTPPA A.A.											
CH 12-958	NMGLKFRQLLWPFHISCLTFFGRETVIEEYLVSFQGVWIRTPPA T.T.G.C.											
CH 6-2625	HMG LKFRQLLWPFHISCLTFFGRETVIEEYLVSFQGVWIRTPPA C.T.C.											
CH 7-1117A	NMGLKFRQLLWPFHISCLTFFGRETVIEEYLVSFQGVWIRTPPA A.A.											
CH 7-2508	NMGLKFRQLLWPFHISCLTFFGRETVIEEYLVSFQGVWIRTPPA A.A.											
CH 9-599	NMGLKFRQLLWPFHISCLTFFGRETVIEEYLVSFQGVWIRTPPA A.A.											
CH 9-7181	NMGLKFRQLLWPFHISCLTFFGRETVIEEYLVSFQGVWIRTPPA T.T.G.C.											
COMPL 11-3028	NMGLKFRQLLWPFHISCLTFFGRETVIEEYLVSFQGVWIRTPPA G.CT.											
COMPL 11-3179	NMGLKFRQLLWPFHISCLTFFGRETVIEEYLVSFQGVWIRTPPA T.T.G.CT.											
COMPL 4-1725	NMGLKFRQLLWPFHISCLTFFGRETVIEEYLVSFQGVWIRTPPA T.A.G.C.											
COMPL 7-2083	NMGLKFRQLLWPFHISCLTFFGRETVIEEYLVSFQGVWIRTPPA A.A.											
X75657	AATATGGGCCCTAAAGTTCAGGCAATTAATGTGGTTTCACATTTCTGTCTCACTTTTGGGAGAGAAAACCGTCATAGAGTATTTGGTGTCTTTCCGAGTGTGGATTCCGACTCCTCCAGCT											
AH 2-10101	NMGLKFRQLLWPFHISCLTFFGRETVIEEYLVSFQGVWIRTPPA .C.C.											

	2300	2310	2320	2330	2340	2350	2360	2370	2380	2390	2400	2410																													
X02763	TATAGACCACCAAATGCCCTATCTTATCAACACTTCGGAACTACTGTTGTTAGACGACGGGACCGAGGCAGGTCGCCCTAGAAGAAGAACTCCCTCGCCTCGCAGACCGAGATCTCCA																																								
AH 2-4890	Y	R	P	P	N	A	P	I	L	S	T	L	P	E	T	T	V	V	R	R	R	D	R	G	R	S	P	R	R	R	T	P	S	P	R	R	R	R	S	P	
AH 1.1	Y	R	P	P	N	A	P	I	L	S	T	L	P	E	T	T	V	V	R	R	R	D	R	G	R	S	P	R	R	R	T	P	S	P	R	R	R	R	S	P	
AH 4-1835	Y	R	P	P	N	A	P	I	L	S	T	L	P	E	T	T	V	V	R	R	R	D	R	G	R	S	P	R	R	R	T	P	S	P	R	R	R	R	S	P	
AH 4-2015	Y	R	P	P	N	A	P	I	L	S	T	L	P	E	T	T	V	V	R	R	R	D	R	G	R	S	P	R	R	R	T	P	S	P	R	R	R	R	S	P	
AH 5-2878	Y	R	P	P	N	A	P	I	L	S	T	L	P	E	T	T	V	V	R	R	R	D	R	G	R	S	P	R	R	R	T	P	S	P	R	R	R	R	S	P	
AH 7-1884	Y	R	P	P	N	A	P	I	L	S	T	L	P	E	T	T	V	V	R	R	R	D	R	G	R	S	P	R	R	R	T	P	S	P	R	R	R	R	S	P	
AH 7-624	Y	R	P	P	N	A	P	I	L	S	T	L	P	E	T	T	V	V	R	R	R	D	R	G	R	S	P	R	R	R	T	P	S	P	R	R	R	R	S	P	
AH 7-640	Y	R	P	P	N	A	P	I	L	S	T	L	P	E	T	T	V	V	R	R	R	D	R	G	R	S	P	R	R	R	T	P	S	P	R	R	R	R	S	P	
AH 8-1993	Y	R	P	P	N	A	P	I	L	S	T	L	P	E	T	T	V	V	R	R	R	D	R	G	R	S	P	R	R	R	T	P	S	P	R	R	R	R	S	P	
AH 10-2400	Y	R	P	P	N	A	P	I	L	S	T	L	P	E	T	T	V	V	R	R	R	D	R	G	R	S	P	R	R	R	T	P	S	P	R	R	R	R	S	P	
AH 10-2596	Y	R	P	P	N	A	P	I	L	S	T	L	P	E	T	T	V	V	R	R	R	D	R	G	R	S	P	R	R	R	T	P	S	P	R	R	R	R	S	P	
AH 11-2641	Y	R	P	P	N	A	P	I	L	S	T	L	P	E	T	T	V	V	R	R	R	D	R	G	R	S	P	R	R	R	T	P	S	P	R	R	R	R	S	P	
AH 12-21	Y	R	P	P	N	A	P	I	L	S	T	L	P	E	T	T	V	V	R	R	R	D	R	G	R	S	P	R	R	R	T	P	S	P	R	R	R	R	S	P	
AH 12-431	Y	R	P	P	N	A	P	I	L	S	T	L	P	E	T	T	V	V	R	R	R	D	R	G	R	S	P	R	R	R	T	P	S	P	R	R	R	R	S	P	
CH 1-6126	Y	R	P	P	N	A	P	I	L	S	T	L	P	E	T	T	V	V	R	R	R	D	R	G	R	S	P	R	R	R	T	P	S	P	R	R	R	R	S	P	
CH 10-1982	Y	R	P	P	N	A	P	I	L	S	T	L	P	E	T	T	V	V	R	R	R	D	R	G	R	S	P	R	R	R	T	P	S	P	R	R	R	R	S	P	
CH 12-2805	Y	R	P	P	N	A	P	I	L	S	T	L	P	E	T	T	V	V	R	R	R	D	R	G	R	S	P	R	R	R	T	P	S	P	R	R	R	R	S	P	
CH 12-82	Y	R	P	P	N	A	P	I	L	S	T	L	P	E	T	T	V	V	R	R	R	D	R	G	R	S	P	R	R	R	T	P	S	P	R	R	R	R	S	P	
CH 5-2628	Y	R	P	P	N	A	P	I	L	S	T	L	P	E	T	T	V	V	R	R	R	D	R	G	R	S	P	R	R	R	T	P	S	P	R	R	R	R	S	P	
CH 6-444A	Y	R	P	P	N	A	P	I	L	S	T	L	P	E	T	T	V	V	R	R	R	D	R	G	R	S	P	R	R	R	T	P	S	P	R	R	R	R	S	P	
COMPL 4-2971	Y	R	P	P	N	A	P	I	L	S	T	L	P	E	T	T	V	V	R	R	R	D	R	G	R	S	P	R	R	R	T	P	S	P	R	R	R	R	S	P	
V01460	TATAGACCACCAAATGCCCTATCTTATCAACACTTCGGAGACTACTGTTGTTAGACGACG--AGGCAGGTCGCCCTAGAAGAAGAACTCCCTCGCCTCGCAGACGAAGTCTCCA																																								
X02496	Y	R	P	P	N	A	P	I	L	S	T	L	P	E	T	T	V	V	R	R	R				G	R	S	P	R	R	R	T	P	S	P	R	R	R	R	S	Q
AH 10-1108	Y	R	P	P	N	A	P	I	L	S	T	L	P	E	T	T	V	V	R	R	R				G	R	S	P	R	R	R	T	P	S	P	R	R	R	R	S	Q
AH 10-1612	Y	R	P	P	N	A	P	I	L	S	T	L	P	E	T	T	V	V	R	R	R				G	R	S	P	R	R	R	T	P	S	P	R	R	R	R	S	Q
CH 10-3329	Y	R	P	P	N	A	P	I	L	S	T	L	P	E	T	T	V	V	R	R	R				G	R	S	P	R	R	R	T	P	S	P	R	R	R	R	S	Q
AH 10-2085	Y	R	P	P	N	A	P	I	L	S	T	L	P	E	T	T	V	V	R	R	R				G	R	S	P	R	R	R	T	P	S	P	R	R	R	R	S	Q
AH 10-2425	Y	R	P	P	N	A	P	I	L	S	T	L	P	E	T	T	V	V	R	R	R				G	R	S	P	R	R	R	T	P	S	P	R	R	R	R	S	Q
AH 10-737	Y	R	P	P	N	A	P	I	L	S	T	L	P	E	T	T	V	V	R	R	R				G	R	S	P	R	R	R	T	P	S	P	R	R	R	R	S	Q
AH 10-486	Y	R	P	P	N	A	P	I	L	S	T	L	P	E	T	T	V	V	R	R	R				G	R	S	P	R	R	R	T	P	S	P	R	R	R	R	S	Q
AH 11-1941	Y	R	P	P	N	A	P	I	L	S	T	L	P	E	T	T	V	V	R	R	R				G	R	S	P	R	R	R	T	P	S	P	R	R	R	R	S	Q
AH 11-2336A	Y	R	P	P	N	A	P	I	L	S	T	L	P	E	T	T	V	V	R	R	R				G	R	S	P	R	R	R	T	P	S	P	R	R	R	R	S	Q
AH 11-2514	Y	R	P	P	N	A	P	I	L	S	T	L	P	E	T	T	V	V	R	R	R				G	R	S	P	R	R	R	T	P	S	P	R	R	R	R	S	Q
AH 11-2524	Y	R	P	P	N	A	P	I	L	S	T	L	P	E	T	T	V	V	R	R	R				G	R	S	P	R	R	R	T	P	S	P	R	R	R	R	S	Q
AH 11-2759	Y	R	P	P	N	A	P	I	L	S	T	L	P	E	T	T	V	V	R	R	R				G	R	S	P	R	R	R	T	P	S	P	R	R	R	R	S	Q
AH 11-2866	Y	R	P	P	N	A	P	I	L	S	T	L	P	E	T	T	V	V	R	R	R				G	R	S	P	R	R	R	T	P	S	P	R	R	R	R	S	Q
AH 11-5664	Y	R	P	P	N	A	P	I	L	S	T	L	P	E	T	T	V	V	R	R	R				G	R	S	P	R	R	R	T	P	S	P	R	R	R	R	S	Q
AH 11-965	Y	R	P	P	N	A	P	I	L	S	T	L	P	E	T	T	V	V	R	R	R				G	R	S	P	R	R	R	T	P	S	P	R	R	R	R	S	Q
AH 12-438	Y	R	P	P	N	A	P	I	L	S	T	L	P	E	T	T	V	V	R	R	R				G	R	S	P	R	R	R	T	P	S	P	R	R	R	R	S	Q
AH 3-577	Y	R	P	P	N	A	P	I	L	S	T	L	P	E	T	T	V	V	R	R	R				G	R	S	P	R	R	R	T	P	S	P	R	R	R	R	S	Q
AH 2-2658	Y	R	P	P	N	A	P	I	L	S	T	L	P	E	T	T	V	V	R	R	R				G	R	S	P	R	R	R	T	P	S	P	R	R	R	R	S	Q
AH 3-7991	Y	R	P	P	N	A	P	I	L	S	T	L	P	E	T	T	V	V	R	R	R				G	R	S	P	R	R	R	T	P	S	P	R	R	R	R	S	Q
AH 11-8035	Y	R	P	P	N	A	P	I	L	S	T	L	P	E	T	T	V	V	R	R	R				G	R	S	P	R	R	R	T	P	S	P	R	R	R	R	S	Q
AH 3-7755	Y	R	P	P	N	A	P	I	L	S	T	L	P	E	T	T	V	V	R	R	R				G	R	S	P	R	R	R	T	P	S	P	R	R	R	R	S	Q
AH 5-920	Y	R	P	P	N	A	P	I	L	S	T	L	P	E	T	T	V	V	R	R	R				G	R	S	P	R	R	R	T	P	S	P	R	R	R	R	S	Q
AH 3-987	Y	R	P	P	N	A	P	I	L	S	T	L	P	E	T	T	V	V	R	R	R				G	R	S	P	R	R	R	T	P	S	P	R	R	R	R	S	Q
AH 4-2307	Y	R	P	P	N	A	P	I	L	S	T	L	P	E	T	T	V	V	R	R	R				G	R	S	P	R	R	R	T	P	S	P	R	R	R	R	S	Q
AH 4-3445	Y	R	P	P	N	A	P	I	L	S	T	L	P	E	T	T	V	V	R	R	R				G	R	S	P	R	R	R	T	P	S	P	R	R	R	R	S	Q
AH 4-775	Y	R	P	P	N	A	P	I	L	S	T	L	P	E	T	T	V	V	R	R	R				G	R	S	P	R	R	R	T	P	S	P	R	R	R	R	S	Q

	2300	2310	2320	2330	2340	2350	2360	2370	2380	2390	2400	2410	
V01460	TATAGACCACCAAATGCCCTATCTTATCAACACTTCGGAGACTACTGTTGTTAGACGACG-----AGGCAGGTCCCTAGAGAAGAACTCCCTCGCCTCGCAGACGAAGTCTCAA												
X02496	Y R P P N A P I L S T L P E T T V V R R R R	G R S P R R R R T P S P R R R R R S Q											
AH 4-778	T												
AH 6.1	T												
AH 5-556	T												
AH 6-1289	T												
AH 6-217	T												
AH 6-220	T												
AH 6-2394	T A												
AH 6-2616	T												
AH 7-1010	T A												
AH 7-1275	T												
AH 7-1440	T												
AH 7-1442	T T A												
AH 7-1501	T A												
AH 7-283	T												
AH 7-3552	T A												
AH 8-1853	T												
AH 8-2186	T												
AH 9-2772	T T T G A A T												
AH 9-296	T												
AH 9-3113	T												
AH 9-3163	T												
AH 9-4070	T												
CH 1-1546	T C												
CH 1-4936	T												
CH 10-3335	T												
CH 11-2343A	T												
CH 11-2866	T												
CH 11-5664	T												
CH 11-884	T												
CH 12-2178	T												
CH 12-958	T												
CH 6-2625	T T G												
CH 7-1117A	T C												
CH 7-2508	T												
CH 9-599	T												
CH 9-7181	T												
COMPL 11-3028	T												
COMPL 11-3179	T												
COMPL 4-1725	T												
COMPL 7-2083	T												
X75657	TATAGACCACCAAATGCCCTATCTTATCAACACTTCGGAGAACTACTGTTGTTAGACGAAG-----AGGCAGGTCCCTAGAGAAGAACTCCCTCGCCTCGCAGACGAAGATCTCAA												
AH 2-10101	T C												

	2420	2430	2440	2450
X02763	TCGCCGCGTCGCAGAGATCTCAATCTCGGGAATCTCAATGT			
AH 2-4890	S P R R R R R S Q S R E S Q C			
AH 1.1	S P R R R R R S Q S R E S Q C			
AH 4-1835	S P R R R R R S Q S R E S Q C			
AH 4-2015	S P R R R R R S Q S R E S Q C			
AH 5-2878	S P R R R R R S Q S R E S Q C			
AH 7-1884	S P R R R R R S Q S R E S Q C			
AH 7-624	S P R R R R R S Q S R E S Q C			
AH 7-640	S P R R R R R S Q S R E S Q C			
AH 8-1993	S P R R R R R S Q S R E S Q C			
AH 10-2400	S P R R R R R S Q S R E S Q C			
AH 10-2596	S P R R R R R S Q S R E S Q C			
AH 11-2641	S P R R R R R S Q S R E S Q C			
AH 12-21	S P R R R R R S Q S R E S Q C			
AH 12-431	S P R R R R R S Q S R E S Q C			
CH 1-6126	S P R R R R R S Q S R E S Q C			
CH 10-1982	S P R R R R R S Q S R E S Q C			
CH 12-2805	S P R R R R R S Q S R E S Q C			
CH 12-82	S P R R R R R S Q S R E S Q C			
CH 5-2628	S P R R R R R S Q S R E S Q C			
CH 6-444A	S P R R R R R S Q S R E S Q C			
COMPL 4-2971	S P R R R R R S Q S R E S Q C			
V01460	TCGCCGCGTCGCAGAGATCTCAATCTCGGGAATCTCAATGT			
X02496	S P R R R R R S Q S R E S Q C			
AH 10-1108	S P R R R R R S Q S R E S Q C			
AH 10-1612	S P R R R R R S Q S R E S Q C			
CH 10-3329	S P R R R R R S Q S R E S Q C			
AH 10-2085	S P R R R R R S Q S R E S Q C			
AH 10-2425	S P R R R R R S Q S R E S Q C			
AH 10-737	S P R R R R R S Q S R E S Q C			
AH 10-486	S P R R R R R S Q S R E S Q C			
AH 11-1941	S P R R R R R S Q S R E S Q C			
AH 11-2336A	S P R R R R R S Q S R E S Q C			
AH 11-2514	S P R R R R R S Q S R E S Q C			
AH 11-2524	S P R R R R R S Q S R E S Q C			
AH 11-2759	S P R R R R R S Q S R E S Q C			
AH 11-2866	S P R R R R R S Q S R E S Q C			
AH 11-5664	S P R R R R R S Q S R E S Q C			
AH 11-965	S P R R R R R S Q S R E S Q C			
AH 12-438	S P R R R R R S Q S R E S Q C			
AH 3-577	S P R R R R R S Q S R E S Q C			
AH 2-2658	S P R R R R R S Q S R E S Q C			
AH 3-7991	S P R R R R R S K S R D S Q C			
AH 11-8035	S P R R R R R S Q S R E S Q C			
AH 3-7755	S P R R R R R S Q S R E S Q C			
AH 5-920	S P R R R R R S Q S R E S Q C			
AH 3-987	S P R R R R R S Q S R E S Q C			
AH 4-2307	S P R R R R R S Q S R E S Q C			
AH 4-3445	S P R R R R R S Q S R E S Q C			
AH 4-775	S P R R R R R S Q S R E S Q C			

	2420	2430	2440	2450
V01460	TCGCCGCGTCGCAGAAGATCTCAATCTCGGGAATCTCAATGT			
X02496	S P R R R R R S Q S R E S Q C			
AH 4-778	S P R R R R R S Q S R E S Q C			
AH 6.I	S P R R R R R S Q S R E S Q C			
AH 5-556	S P R R R R R S Q S R E S Q C			
AH 6-1289	S P R R R R R S Q S R E S Q C			
AH 6-217	S P R R R R R S Q S R E S Q C			
AH 6-220	S P R R R R R S Q S R E S Q C			
AH 6-2394	S P R R R R R S Q S R E S Q C			
AH 6-2616	S P R R R R R S Q S R E S Q C			
AH 7-1010	S P R R R R R S Q S R E S Q C			
AH 7-1275	S P R R R R R S Q S R E S Q C			
AH 7-1440	S P R R R R R S Q S R E S Q C			
AH 7-1442	S P R R R R R S Q S R E S Q C			
AH 7-1501	S P R R R R R S Q S R E S Q C			
AH 7-283	S P R R R R R S Q S R E S Q C			
AH 7-3552	S P R R R R R S Q S R E S Q C			
AH 8-1853	S P R R R R R S Q S R E S Q C			
AH 8-2186	S P R R R R R S Q S R E S Q C			
AH 9-2772	S P R R R R R S Q S R E S Q C			
AH 9-296	S P R R R R R S Q S R E S Q C			
AH 9-3113	S P R R R R R S Q S R E S Q C			
AH 9-3163	S P R R R R R S Q S R E S Q C			
AH 9-4070	S P R R R R R S Q S R E S Q C			
CH 1-1546	S P R R R R R S Q S R E S Q C			
CH 1-4936	S P R R R R R S Q S R E S Q C			
CH 10-3335	S P R R R R R S Q S R E S Q C			
CH 11-2343A	S P R R R R R S Q S R E S Q C			
CH 11-2866	S P R R R R R S Q S R E S Q C			
CH 11-5664	S P R R R R R S Q S R E S Q C			
CH 11-884	S P R R R R R S Q S R E S Q C			
CH 12-2178	S P R R R R R S Q S R E S Q C			
CH 12-958	S P R R R R R S Q S R E S Q C			
CH 6-2625	S P R R R R R S Q S R E S Q C			
CH 7-1117A	S P R R R R R S Q S R E S Q C			
CH 7-2508	S P R R R R R S Q S R E S Q C			
CH 9-599	S P R R R R R S Q S R E S Q C			
CH 9-7181	S P R R R R R S Q S R E S Q C			
COMPL 11-3028	S P R R R R R S Q S R E S Q C			
COMPL 11-3179	S P R R R G R S Q S R E S Q C			
COMPL 4-1725	S P R R R R R S Q S R E S Q C			
COMPL 7-2083	S P R R R R R S Q S R E S Q C			
X75657	TCGCCGCGTCGCAGAAGATCTCAATCTCCAGCTCCCAATGT			
AH 2-10101	S P R R R R R S Q S P A S Q C			

Appendix Part B. Sequences of HBV region preS/S obtained during the study.

	2860	2870	2880	2890	2900	2910	2920	2930	2940	2950	2960	2970
X02763	ATGGGAGGTTGGTCATCAAAACCTCGCAAAGGCATGGGGACGAATCTTCTGTTCCTCAATCCTCTGGGATTCTTTCCCGATCATCAGTTGGACCTTCATTTCGGAGCCAACCTCAAAACAAT											
	M G G W S S K P R K G M G T N L S V P N P L G F F P D H Q L D P A F G A N S N N											
AH 1. II	-----											
AH 4. I	-----											
AH 4-1835	-----C-----											
AH 4-2015	M G G W S S K P R K G M G T N L S V P N P L G F F P D H Q L D P A F G A N S N N											
AH 5-2878	-----C-----											
AH 7-624	-----C-----											
AH 7-640	-----C-----											
AH 9-454	-----											
AH 7-1884	-----											
AH 10-2400	-----											
AH 10-2596	-----											
CH 3-3608	-----C-----											
CH 5-2628	-----C-----											
CH 6-444A	-----C-----											
CH 11-913	-----											
CH 12-82	-----											
CH 12-2805	-----											
COMPL 4-2971	-----C-----											
COMPL 9-454	M G G W S S K P R K G M G T N L S V P N P L G F F P D H Q L D P A F G A N S N N											
COMPL 9-1553	-----											
V01460	-----ATGGGGCAGAAATCTTCCACCAGCAATCCTCTGGGATTCTTTCCCGACCACCAAGTTGGATCCAGCCTTCAGAGCAAACACCCGAAAT											
	M G Q N L S T S N P L G F F P D H Q L D P A F R A N T A N											
X02496	-----M G Q N L S T S N P L G F F P D H Q L D P A F R A N T A N											
AH 2-2658	-----											
AH 3. I	-----											
AH 3-577	-----CAG-----											
AH 3-987	M G R N L S T S N P L G F F P D H Q L D P A F R A N T A N											
AH 3-7755	M G Q N L S T S N P L G F F P D H Q L D P A F R A N T A N											
AH 3-7991	M G Q N L S T S N P L G F F P D H Q L D P A F R A N T A N											
AH 4-775	M G Q N L S T S N P L G F F P D H Q L D P A F R A N T A N											
AH 4-778	M G Q N L S T S N P L G F F P D H Q L D P A F R A N T A N											
AH 4-2307	M G Q N L S T S N P L G F F P D H Q L D P A F R A N T A N											
AH 4-3445	M G Q N L S T S N P L G F F P D H Q L D P A F R A N T A N											
AH 5. I	-----											
AH 5-556	-----											
AH 5-920	-----A-----											
AH 5-1976	M G Q N L S T S N P L G F F P D H Q L D P A F R A N T A N											
AH 6. I	-----											
AH 6-217	-----A-----											
AH 6-220	M G Q N L S T S N P L G F F P D H Q L D P A F R A N T A N											
AH 7-283	-----											
	M G Q N L S T S N P L G F F P D H Q L D P A F R A N T A N											

	2860	2870	2880	2890	2900	2910	2920	2930	2940	2950	2960	2970
V01460ATGGGGCAGAATCTTTCCACCAGCAATCCTCTGGGATTTTCCCGACCACCAGTTGGATCCAGCCTTCAGAGCAAACACCGAAAT.....											
X02496M G Q N L S T S N P L G F F P D H Q L D P A F R A N T A N.....											
AH 7-496M G Q N L S T S N P L G F F P D H Q L D P A F R A N T A N.....											
AH 7-1010M G Q N L S T S N P L G F F P D H Q L D P A F R A N T A N.....											
AH 7-1275											
AH 7-1442											
AH 8-2186											
AH 9-2772											
AH 9-296											
AH 9-297											
AH 9-3163											
AH 9-3113											
AH 9-3897											
AH 9-4070											
AH 10-486											
AH 10-737											
AH 10-885											
AH 10-1108											
AH 10-1612											
AH 10-2085											
AH 10-2425											
AH 11-965											
AH 11-1941											
AH 11-2165											
AH 11-2514											
AH 11-2759											
AH 12-116M G Q N L S T S N P L G F F P D H Q L D P A F R A N T A N.....											
CH 1-1546M G Q N L S T S N P L G F F P D H Q L D P A F R A N T T N.....											
CH 6-2107											
CH 7-1117AM G Q N L S T S N P L G F F P D H Q L D P A F R A N T A N.....											
CH 9-599M G Q N L S T S N P L G F F P D H Q L D P A F R A N T A N.....											
CH 11-884M G Q N L S T S N P L G F F P D H Q L D P A F R A N T A N.....											
CH 11-2343A											
CH 12-958											
COMPL 4-1725M G Q N L S T S N P L G F F P D H Q L D P A F R A N T A N.....											
COMPL 11-302											
COMPL 11-317M G Q N L S T S N P L G F F P D H Q L D P A F R A N T A N.....											
X75657ATGGGGAAGAATATTTCCACCACCAATCCTCTGGGATTTTCCCGACCACCAGTTGGATCCAGCATTTCAGAGCAAACACCGAAAT.....											
AH 2-10101M G K N I S T T N P L G F F P D H Q L D P A F R A N T R N.....											

	2980	2990	3000	3010	3020	3030	3040	3050	3060	3070	3080	3090
X02763	CCAGATTGGGACTTCAACCCCGTCAAGGACGACTGGCCAGCAGCCCAACCAAGTAGGAGTGGGAGCATTGCGGCCAAGGCTCACCCCTCCACACGGCGGTATTTGGGGTGGAGCCCTCAG											
AH 1. II	P D W D D F N P V K D D W P A A N Q V G V G A F G P R L T P P H G G I L G W S P Q											
AH 4. I	-----											
AH 4-1835A.....TC.....A..G.....G.....											
AH 4-2015	P D W D D F N P I K D H W P A A K Q V G V G A F G P G L T P P H G G I L G W S P Q											
AH 5-2878A.....C.....G.....G.....											
AH 7-624	P D W D D F N P I K D H W P A A N Q V G V G A F G P G L T P P H G G I L G W S P Q											
AH 7-640A.....C.....G.....G.....											
AH 9-454	P D W D D F N P V K D D W P A A N Q V G V G A F G P R L T P P H G G I L G W S P Q											
AH 7-1884	-----											
AH 10-2400	-----											
AH 10-2596	-----											
CH 3-3608A.....C.....G.....G.....											
CH 5-2628	P D W D D F N P I K D H W P A A N Q V G V G A F G P G L T P P H G G I L G W S P Q											
CH 6-444AA.....C.....G.....G.....											
CH 11-913	P D W D D F N P I K D H W P A A N Q V G V G A F G P G L T P P H G G I L G W S P Q											
CH 12-82	-----											
CH 12-2805	-----											
COMPL 4-2971A.....T.....											
COMPL 9-454	P D W D D F N P I K D D W P A A N Q V G V G A F G P R L T P P H G G I L G W S P Q											
COMPL 9-1553	-----											
V01460	CCAGATTGGGACTTCAATCCCAACAAGGACACCTGGCCAGCAGCCCAACAAGTAGGAGCTGGAGCATTGCGGCTGGGTTTACCCACCAGCAGGAGCCTTTGGGGTGGAGCCCTCAG											
X02496	P D W D D F N P N K D T W P D A N K V G A G A F G L G F T P P H G G L L G W S P Q											
AH 2-2658	P D W D D F N P N K D T W P D A N K V G A G A F G L G F T P P H G G L L G W S P Q											
AH 3. I	-----											
AH 3-577T.....A.....A.....											
AH 3-987	P D W D D F N P N K D T W P D A N K V G A G A F G L G F T P P H G G L L G W S P Q											
AH 3-7755A.....A.....											
AH 3-7991	P D W D D F N P N K D T W P D A N K V G A G A F G L G F T P P H G G L L G W S P Q											
AH 4-775T.....A.....A.....											
AH 4-778	P D W D D F N P N K D T W P D A N K V G A G A F G L G F T P P H G G L L G W S P Q											
AH 4-2307A.....C.....											
AH 4-3445	P D W D D F N P N K D T W P D A N K V G A G A F G L G F T P P H G G L L G W S P Q											
AH 5. I	-----											
AH 5-556	-----											
AH 5-920A.....											
AH 5-1976	P D W D D F N P N K D T W P D A N K V G A G A F G L G F T P P H G G L L G W S P Q											
AH 6. I	-----											
AH 6-217A.....											
AH 6-220	P D W D D F N P N K D T W P D A N K V G A G A F G L G F T P P H G G L L G W S P Q											
AH 7-283C.....											
	P D W D D F N P N K D T W P D A N K V G A G A F G P G F T P P H G G L L G W S P Q											

	2980	2990	3000	3010	3020	3030	3040	3050	3060	3070	3080	3090
V01460	CCAGATTGGGACTTCAATCCCAACAAGGACACCTGGCCAGACGCCAACAAGGTAGGAGCTGGAGCATTGGGGCTGGGTTTCAACCCACCGCACGGAGGCCTTTGGGGTGGAGCCCTCAG											
X02496	P D W D D F N P N K D T W P D A N K V G A G A F G L G F T P P H G G L L G W S P Q											
AH 7-496	P D W D D F N P N K D T W P D A N K V G A G A F G L G F T P P H G G L L G W S P Q											
AH 7-1010	P D W D D F N P N K D T W P D A N K V G A G A F G L G F T P P H G G L L G W S P Q											
AH 7-1275	-----											
AH 7-1442	-----											
AH 8-2186	-----											
AH 9-2772	-----											
AH 9-296	-----											
AH 9-297	-----											
AH 9-3163	-----											
AH 9-3113	-----											
AH 9-3897	-----											
AH 9-4070	-----											
AH 10-486	-----											
AH 10-737	-----											
AH 10-885	-----											
AH 10-1108	-----											
AH 10-1612	-----											
AH 10-2085	-----											
AH 10-2425	-----											
AH 11-965	-----											
AH 11-1941	-----											
AH 11-2165	-----											
AH 11-2514	-----											
AH 11-2759	-----											
AH 12-116	P D W D D F N P N K D T W P D A N K V G A G A F G L G F T P P H G G L L G W S P Q											
CH 1-1546	P D W D D F N P N K D P W P D A N K V G A G A F G P G F T P P H G G L L G W S P Q											
CH 6-2107	-----											
CH 7-1117A	P D W D D F N P N K D T W P D A N K V G A G A F G P G F T P P H G G L L G W S P Q											
CH 9-599	P D W D D F N P N K D T W P D A N K V G A G A F G L G F T P P H G G L L G W S P Q											
CH 11-884	-----											
CH 11-2343A	-----											
CH 12-958	-----											
COMPL 4-1725	P D W D D F N P N K D T W P D A N K V G A G A F G L G F T P P H G G L L G W S P Q											
COMPL 11-3028	-----											
COMPL 11-3179	P D W D D F N P N K D T W P D A N K V G A G A F G L G F T P P H G G L L G W S P Q											
X75657	CCAGATTGGGACCACAAATCCCAACAAGGACCACTGGACAGAGCCAACAAGGTAGGAGTGGGAGCATTGGGGCTGGGTTCACTCCCCACCGCACGGAGGCCTTTGGGGTGGAGCCCTCAG											
AH 2-10101	P D W D D H N P N K D H W T E A N K V G V G A F G P G L T P P H G G L L G W S P Q											

3100 3110 3120 3130 3140 3150 3160 3170 3180 3190 3200 3210

X02763 GCTCAGGGCATATTGACCCAGTGTCAACAATTCCTCCTCGCTCCACCAATCGGCAGTCAGGAAGGCAGCCTACTCCCATCTCTCCACCTTAAGAGACAGTCATCCTCAGGCCATG
A Q G I L T T V S T I P P P A S T N R Q S G R Q P T P I S P P L R D S H P Q A M

AH 1. II -----

AH 4. I -----

AH 4-1835 -----

AH 4-2015 -----

AH 5-2878 -----

AH 7-624 -----

AH 7-640 -----

AH 9-454 -----

AH 7-1884 -----

AH 10-2400 -----

AH 10-2596 -----

CH 3-3608 -----

CH 5-2628 -----

CH 6-444A -----

CH 11-913 -----

CH 12-82 -----

CH 12-2805 -----

COMPL 4-2971 -----

COMPL 9-454 -----

COMPL 9-1553 -----

V01460 GCTCAGGGCATACTACAAACTTTGCCAGCAAAATCCGCCTCCTCGCTCCACCAATCGGCAGTCAGGAAGGCAGCCTACCCCGCTGTCTCCACCTTTGAGAAACACTCATCCTCAGGCCATG
A Q G I L Q T L P A N P P P A S T N R Q S G R Q P T P L S P P L R N T H P Q A M

X02496 A Q G I L G G T L P A N P P P A S T N R Q S G R Q P T P L S P P L R N T H P Q A M
A Q G I L E T L P A N P P P A S T N R Q S G R Q P T P L S P P L R N T H P Q A M

AH 2-2658 -----

AH 3. I -----

AH 3-577 -----

AH 3-987 -----

AH 3-7755 -----

AH 3-7991 -----

AH 4-775 -----

AH 4-778 -----

AH 4-2307 -----

AH 4-3445 -----

AH 5. I -----

AH 5-556 -----

AH 5-920 -----

AH 5-1976 -----

AH 6. I -----

AH 6-217 -----

AH 6-220 -----

AH 7-283 -----

	3100	3110	3120	3130	3140	3150	3160	3170	3180	3190	3200	3210
V01460	GCTCAGGGCATACTACAAACTTTGCCAGCAAATCCGCCTCCTGCCTCCACCAATCGCCAGTCAGGAAGGCAGCCTACCCCGCTGTCTCCACCTTTGAGAAACACTCATCCTCAGGCCATG											
X02496	A Q G I L Q T L P A N P P P A S T N R Q S G R Q P T P L S P P L R N T H P Q A M											
AH 7-496	A Q G I L E T L P A N P P P A S T N R Q S G R Q P T P L S P P L R N T H P Q A M											
AH 7-1010	A Q G I I Q T L P A N P P P A S T N R Q S G R Q P T P L S P P L R N T H P Q A M											
AH 7-1275	-----											
AH 7-1442	-----											
AH 8-2186	-----											
AH 9-2772	-----											
AH 9-296	-----											
AH 9-297	-----											
AH 9-3163	-----											
AH 9-3113	-----											
AH 9-3897	-----											
AH 9-4070	-----											
AH 10-486	-----											
AH 10-737	-----											
AH 10-885	-----											
AH 10-1108	-----											
AH 10-1612	-----											
AH 10-2085	-----											
AH 10-2425	-----											
AH 11-965	-----											
AH 11-1941	-----											
AH 11-2165	-----											
AH 11-2514	-----											
AH 11-2759	-----											
AH 12-116T.....A.....C.....G.....A..T.....											
CH 1-1546	A H G I I Q T L P A N P P P A S T N R Q S G R Q P T P L S P P L R N T H P Q A M											
CH 6-2107	A Q G I L Q T L P A N P P P A S S N R Q S G R Q P T P L S P P L R N T H P Q A M											
CH 7-1117AG...G.....T.....											
CH 9-599	A Q G I L E T L P A N P P P A S T N R Q S G R Q P T P L S P P L R N T H P Q A M											
CH 11-884	A Q G I I Q T L P A N P P P A S T N R Q S G R Q P T P L S P P L R N T H P Q A M											
CH 11-2343A	-----											
CH 12-958	-----											
COMPL 4-1725A.....C.....A..T..C.....G.....											
COMPL 11-3028	A Q G I I Q T L P A N P P P A S T H R Q S G R Q P T P L S P P L R N T H P Q A M											
COMPL 11-3179A.....C.....A..T.....A.....											
X75657	A Q G I I Q T L P A N P P P A S T N R Q S G R Q P T P L S P P L R N T H P Q A M											
AH 2-10101	GCTCAAGGCATGCTAAAAACATTGCCAGCAGATCCGCCTCCTGCCTCCACCAATCGCCAGTCAGGAAGGCAGCCTACCCCAATCACTCCACCTTTGAGAGACACTCATCCTCAGGCCATG											
	A Q G M L K T L P A D P P P A S T N R Q S G R Q P T P I T P P L R D T H P Q A M											
	A Q G M L K T L P A D P P P A S T N R Q S G R Q P T P I T P P L R D T H P Q A M											

3220 10 20 30 40 50 60 70 80 90 100 110

X02763 CAGTGGAAATCCACTGCCTTCCACCAAACTCTGCAAGATCCCAGAGTCAGGGTCTGTATCTTCCCTGCTGGTGGCTCCAGTTCAGGAACAGTAAACCCCTGCTCCGAATATTGCCTCTCAC
 Q W N S T A F H Q T L Q D P R V R G L Y L P A G G S S S S G T V N P A P N I A S H

AH 1. II -----

AH 4. I -----

AH 4-1835G.....T.....
 Q W N S T A F H Q A L Q D P R V R G L Y F P A G G S S S S G T V N P A P N I A S H

AH 4-2015 -----

AH 5-2878G.....T.....
 Q W N S T A F H Q A L Q D P R V R G L Y F P A G G S S S S G T V N P A P N I A S H

AH 7-624G.....T.....
 Q W N S T A F H Q A L Q D P R V R G L Y F P A G G S S S S G T V N P A P N I A S H

AH 7-640G.....T.....
 Q W N S T A F H Q A L Q D P R V R G L Y F P A G G S S S S G T V N P A P N I A S H

AH 9-454 -----

AH 7-1884 -----

AH 10-2400 -----

AH 10-2596 -----

CH 3-3608G.....T.....
 Q W N S T A F H Q A L Q D P R V R G L Y F P A G G S S S S G T V N P A P N I A S H

CH 5-2628G.....T.....
 Q W N S T A F H Q A L Q D P R V R G L Y F P A G G S S S S G T V N P A P N I A S H

CH 6-444AG.....T.....
 Q W N S T A F H Q A L Q D P R V R G L Y F P A G G S S S S G T V N P A P N I A S H

CH 11-913 -----

CH 12-82 -----

CH 12-2805 -----

COMPL 4-2971G.....T.....
 Q W N S T A F H Q A L Q D P R V R G L Y F P A G G S S S S G T V N P A P N I A S H

COMPL 9-454 -----

COMPL 9-1553 -----

V01460 CAGTGGAAATCCACAACCTTCCACCAAACTCTGCAAGATCCCAGAGTCAGAGGCCTGTATTTCCCTGCTGGTGGCTCCAGTTCAGGAACAGTAAACCCCTGTTCTGACTACTGCCTCTCCC
 Q W N S T T F H Q T L Q D P R V R G L Y F P A G G S S S S G T V N P V L T T A S P

X02496C.....T.....
 Q W N S T T F H Q T L Q D P R V R G L Y F P A G G S S S S G T V N P V P T T V S P

AH 2-2658 -----

AH 3. I -----

AH 3-577C.....G.....C.....
 Q W N S T T F H Q T L Q D P R V R G L Y F P A G G S S S S G T V N P V P T T A S P

AH 3-987C.....G.....C.....
 Q W N S T T F H Q T L Q D P R V R G L Y F P A G G S S S S G T V N P V P T T A S P

AH 3-7755C.....G.....C.....G.....C.....
 Q W N S T T F H Q T L Q D P R V R G L Y L P A G G S S S S G T V N P V P T T A S P

AH 3-7991T.....C.....A.....
 Q W N S T T F H Q T L Q D P R V R G L Y F P A G G S S S S G T V N P V P T T A S H

AH 4-775C.....G.....C.....
 Q W N S T T F H Q T L Q D P R V R G L Y F P A G G S S S S G T V N P V P T T A S P

AH 4-778C.....G.....C.....
 Q W N S T T F H Q T L Q D P R V R G L Y F P A G G S S S S G T V N P V P T T A S P

AH 4-2307C.....G.....C.....
 Q W N S T T F H Q T L Q D P R V R G L Y F P A G G S S S S G T V N P V P T T A S P

AH 4-3445C.....C.....A.....
 Q W N S T T F H Q T L Q D P R V R G L Y F P A G G S S S S G T V N P V P T T A S H

AH 5. I -----

AH 5-556 -----

AH 5-920C.....G.....C.....
 Q W N S T T F H Q T L Q D P R V R G L Y F P A G G S S S S G T V N P V P T T A S P

AH 5-1976 -----

AH 6. I -----

AH 6-217C.....T.....
 Q W N S T T F H Q T L Q D P R V R G L Y F P A G G S S S S G T V N P V P T T V S P

AH 6-220G.....C.....
 T T F H Q T L Q D P R V R G L Y F P A G G S S S S G T V N P V P T T A S P

AH 7-283C.....T.....G.....A.....
 Q W N S T T F H Q T L Q D P R V R G L Y F P A G G S S S S G T V N P V P T T A S H

	3220	10	20	30	40	50	60	70	80	90	100	110
V01460	CAGTGGAAATCCACAACCTCCACCAAACTCTGCAAGATCCCAGAGTGAGAGGCCTGTATTTCCCTGCTGGTGGCTCCAGTTCAGGAACAGTAAACCCGTCTGACTACTGCCTCTCCC											
X02496	Q W N S T T F H Q T L Q D P R V R G L Y F P A G G S S S S G T V N P V L T T A S P											
AH 7-496C.....T.....G.....C.....T.....A.....											
AH 7-1010G.....C.....											
AH 7-1275	-----											
AH 7-1442	-----											
AH 8-2186	-----											
AH 9-2772	-----											
AH 9-296	-----											
AH 9-297	-----											
AH 9-3163	-----											
AH 9-3113	-----											
AH 9-3897	-----											
AH 9-4070	-----											
AH 10-486	-----											
AH 10-737	-----											
AH 10-885	-----											
AH 10-1108	-----											
AH 10-1612	-----											
AH 10-2085	-----											
AH 10-2425	-----											
AH 11-965	-----											
AH 11-1941	-----											
AH 11-2165	-----											
AH 11-2514	-----											
AH 11-2759	-----											
AH 12-116C.....G.....C.....											
CH 1-1546C.....T.....C.....C.....A.....											
CH 6-2107	-----											
CH 7-1117AC.....T.....G.....C.....A.....											
CH 9-599C.....C.....T.....A.....											
CH 11-884	-----											
CH 11-2343A	-----											
CH 12-958	-----											
COMPL 4-1725C.T.....G.....T.....C.....											
COMPL 11-3028	Q W N S T T F H Q T L Q D P R V R G L Y F P A G G S S S S G I V N P V P T T A S P											
COMPL 11-3179C.....T.....G.....C.....											
X75657	CAGTGGAAATCCACAACATCCACCAAGCTCTGCAGGATCCCAGAGTAAGAGGCCTGTATTTCCCTGCTGGTGGCTCCAGTTCGGGAACAGTGAACCCGTCTCCGACTACTGCCTCACTC											
AH 2-10101	Q W N S T T F H Q A L Q D P R V R G L Y F P A G G S S S S G T V N P V P T T A S L											

	120	130	140	150	160	170	180	190	200	210	220	230
X02763	ATCTCGTCAATCTCCGCGAGGACTGGGGACCTGTGACGAACATGGAGAACATCACATCAGGATTCCTAGGACCCCTGCTCGTGTACAGCGGGGTTTTCTTGTGACAAGAATCCTC											
AH 1. II	I S S I S A R T G D P V T N M E N I T S G F L G P L L V L Q A G F F L L T R I L											L L T R I L
AH 4. I	-----											L L T R I L
AH 4-1835	-----											L L T R I L
AH 4-2015	I S S I S A R T G D P V T N M E N I T S G F L G P L L V L Q A G F F L L T R I L											L L T R I L
AH 5-2878	-----											L L T R I L
AH 7-624	I S S I S A R T G D P V T N M E N I T S G F L G P L L V L Q A G F F L L T R I L											L L T R I L
AH 7-640	I S S I S A R T G D P V A N M E N I T S G F L G P L L V L Q A G F F L L T R I L											L L T R I L
AH 9-454	I S S I S A R T G D P V T N M E N I T S G F L G P L L V L Q A G F F L L T R I L											L L T R I L
AH 7-1884	-----											L L T R I L
AH 10-2400	-----											L L T R I L
AH 10-2596	-----											L L T R I L
CH 3-3608	-----											L L T R I L
CH 5-2628	I S S I S A R T G D P V T N M E N I T S G L L G P L L V L Q A G F F L L T R I L											L L T R I L
CH 6-444A	I S S I S A R T G D P V T N M E N I T S G F L G P L L V L Q A G F F L L T R I L											L L T R I L
CH 11-913	I S S I S A R T G D P V T N M E N I T S G F L G P L L V L Q A G F F L L T R I L											L L T R I L
CH 12-82	-----											L L T R I L
CH 12-2805	-----											L L T R I L
COMPL 4-2971	-----											L L T R I L
COMPL 9-454	I S S I S A R T G D P V T N M E N I T S G F L G P L L V L Q A G F F L L T R I L											L L T R I L
COMPL 9-1553	-----											L L T R I L
V01460	TTATCGTCAATCTTCTCGAGGATTGGGGACCTGCGCTGAACATGGAGAACATCACATCAGGATTCCTAGGACCCCTTCTCGTGTACAGCGGGGTTTTCTTGTGACAAGAATCCTC											
X02496	L S S I F S R I G D P A L N M E N I T S G F L G P L L V L Q A G F F L L T R I L											L L T R I L
AH 2-2658	I S S I F S R I G D P A L N M E N I T S G F L G P L L V L Q A G F F L L T R I L											L L T R I L
AH 3. I	-----											L L T R I L
AH 3-577	A C G											L L T R I L
AH 3-987	I S S I F S R I G D P A L N M E T I T S G F L G P L L V L Q A G F F L L T R I L											L L T R I L
AH 3-7755	I S S I F S R I G D P V L N M E N I T S G F L G P L L V L Q A G F F L L T R I L											L L T R I L
AH 3-7991	I S S I F S R I G D P V L N M E N I T S G F L G P L L V L Q A G F F L L T K I L											L L T R I L
AH 4-775	L S S I F S R I G D P A L N M E N I T S G F L G P L L V L Q A G F F L L T K I L											L L T R I L
AH 4-778	A G											L L T R I L
AH 4-2307	I S S I F S R I G D P A L N M E N I T S G F L G P L L V L Q A G F F L L T R I L											L L T R I L
AH 4-3445	I S S I F S R I G D P V L N M E N I T S G F L G P L L V L Q A G F F L L T R I L											L L T R I L
AH 5. I	L S S I F S R I G D P A L N M E N I T S G F L G P L L V L Q A G F F L L T R I L											L L T R I L
AH 5-556	-----											L L T R I L
AH 5-920	A C G											L L T R I L
AH 5-1976	I S S I F S R I G D P A L N M E N I T S G L L G P L L V L Q A G F F L L T R I L											L L T R I L
AH 6. I	-----											L L T R I L
AH 6-217	A G											L L T R I L
AH 6-220	I S S I F S R I G D P A L N M E N I T S G F L G P L L V L Q A G F F L L T R I L											L L T R I L
AH 7-283	I S S I F S R I G D P V L N M E N I T S G F L G P L L V L Q A G F F L L T R I L											L L T R I L

	120	130	140	150	160	170	180	190	200	210	220	230
V01460	TTATCGTCAATCTTCTCGAGGATTGGGGACCTGCGCTGAACATGGAGAACATCACATCAGGATTCCTAGGACCCCTTCTCGTGTACAGGCGGGTTTTCTTGTGACAAGAATCCTC											
X02496	L S S I F S R I G D P A L N M E N I T S G F L G P L L V L Q A G F F L L T R I L											
AH 7-496	I S S I F S R I G D P A L N M E N I T S G F L G P L L V L Q A G F F L L T R I L											
AH 7-1010	I S S I F S R I G D P A L N M E N I T S G F L G P L L V L Q A G F F L L T R I L											
AH 7-1275	-----											
AH 7-1442	-----											
AH 8-2186	-----											
AH 9-2772	-----											
AH 9-296	-----											
AH 9-297	-----											
AH 9-3163	-----											
AH 9-3113	-----											
AH 9-3897	-----											
AH 9-4070	-----											
AH 10-486	-----											
AH 10-737	-----											
AH 10-885	-----											
AH 10-1108	-----											
AH 10-1612	-----											
AH 10-2085	-----											
AH 10-2425	-----											
AH 11-965	-----											
AH 11-1941	-----											
AH 11-2165	-----											
AH 11-2514	-----											
AH 11-2759	-----											
AH 12-116	A.....T.A.....G.....A.....											
CH 1-1546	I S S I F S R I G D P V M N M E N I T S G F L G P L L V L Q A G F F L L T K I L											
CH 6-2107	L S S I F S R I G D P A L N M E N I T S G F L G P L L V L Q A G F F L L T R I L											
CH 7-1117A	A.....G.....											
CH 9-599	I S S I F S R I G D P A L N M E N I T S G F L G P L L V L Q A G F F L L T R I L											
CH 11-884	L S S I F S R I G D P A L N M E N I T S G F L G P L L V L Q A G F F L L T R I L											
CH 11-2343A	-----											
CH 12-958	-----											
COMPL 4-1725	A.....G.....											
COMPL 11-3028	I S S I F S R I G D P A L N M E N I T S G F L G P L L V L Q A G F F L L T R I L											
COMPL 11-3179	A.....C.....G.....											
X75657	ATTCGTCAATCTTCTCGAGGATTGGGGACCTGCACCGAACATGGAAGCATCACATCAGGATTCCTAGGACCCCTGCTCGTGTACAGGCGGGTTTTCTTGTGACAAAATCCTC											
AH 2-10101	I S S I F S R I G D P A P N M E S I T S G F L G P L L V L Q A G F F L L T K I L											
	I S S I F S R I G D P A P N M E N I T S G F L G P L L V L Q A G F F L L T K I L											

	240	250	260	270	280	290	300	310	320	330	340	350
X02763	ACAATACCGCAGAGTCTAGACTCGTGGTGGACTTCTCTCAATTTTCTAGGGGATCTCCCGTGTCTTGGCCAAAATTCGCAGTCCCCAACCTCCAATCACTACCAACCTCTGTCCCT											
AH 1. II	T I P Q S L D S W W T S L N F L G G S P V C L G Q N S Q S P T S N H S P T S C P	A.										
AH 4. I	T I P Q S L D S W W T S L N F L G G S P V C L G Q N S Q S P T S N H S P T S C P	A.										
AH 4-1835	T I P Q S L D S W W T S L N F L G G S P V C L G Q N S Q S P T S N H S P T S C P	A.										
AH 4-2015	T I P Q S L D S W W T S L N F L G G S P V C L G Q N S Q S P T S N H S L T S C P	T.										
AH 5-2878	T I P Q S L D S W W T S L N F L G G S P V C L G Q N S Q S P T S N H S P T S C P	A.										
AH 7-624	T I P Q S L D S W W T S L N F L G G S P V C L G Q N S Q S P T S N H S P T S C P	A.										
AH 7-640	T I P Q S L D S W W T S L N F L G G S P V C R G Q N S Q S P T S N H S P T S C P	G.										
AH 9-454	T I P Q S L D S W W T S L S F L G G S P V C L G Q N S Q S P T S N H S L T S C P	G. A. T.										
AH 7-1884	T I P Q S L D S W W T S L N F L G G S P V C L G Q N S Q S P T S N H S P T S C P	A.										
AH 10-2400	T I P Q S L D S W W T S L N F L G G S P V C L G Q N S Q S P T S N H S P T S C P	A.										
AH 10-2596	T I P Q S L D S W W T S L N F L G G S P V C L G Q N S Q S P T S N H S P T S C P	A.										
CH 3-3608	T I P Q S L D S W W T S L N F L G G S P V C L G Q N S Q S P T S N H S P T S C P	A.										
CH 5-2628	T I P Q S L D S W W T S L N F L G G S P V C L G Q N S Q S P T S N H S P T S C P	A.										
CH 6-444A	T I P Q S L D S W W T S L N F L G G S P V C L G Q N S Q S P T S N H S P T S C P	A.										
CH 11-913	T I P Q S L D S W W T S L N F L G G S P V C L G Q N S Q S P T S N H S P T S C P	A.										
CH 12-82	T I P Q S L D S W W T S L N F L G G S P V C L G Q N S Q S P T S N H S P T S C P	A.										
CH 12-2805	T I P Q S L D S W W T S L N F L G G S P V C L G Q N S Q S P T S N H S P T S C P	A.										
COMPL 4-2971	T I P Q S L D S W W T S L N F L G G S P V C L G Q N S Q S P T S N H S P T S C P	A.										
COMPL 9-454	T I P Q S L D S W W T S L S F L G G S P V C L G Q N S Q S P T S N H S L T S C P	G. A. T.										
COMPL 9-1553	T I P Q S L D S W W T S L N F L G G S P V C L G Q N S Q S P T S N H S P T S C P	A.										
V01460	ACAATACCGCAGAGTCTAGACTCGTGGTGGACTTCTCTCAATTTTCTAGGGGAACTACCGTGTCTTGGCCAAAATTCGCAGTCCCCAACCTCCAATCACTACCAACCTCTGTCCCT											
X02496	T I P Q S L D S W W T S L N F L G G T T V C L G Q N S Q S P T S N H S P T S C P	C.										
AH 2-2658	T I P Q S L D S W W T S L N F L G G T T V C L G Q N S Q S P T S N H S P T S C P	C.										
AH 3. I	T I P Q S L D S W W T S L N F L G G T T V C L G Q N S Q S P T S N H S P T S C P	C.										
AH 3-577	T I P Q S L D S W W T S L N F L G G T T V C L G Q N S Q S P T S N H S P T S C P	C.										
AH 3-987	T I P Q S L D S W W T S L N F L G G T T V C L G Q N S Q S P T S N H S P T S C P	C.										
AH 3-7755	T I P Q S L D S W W T S L N F L G G T T V C L G Q N S Q S P T S N H S P T S C P	C.										
AH 3-7991	T I P Q N L D S W W T S L S F L G E H T V C L G Q N S Q C P T S N H S P T S C P	A. G. A. CAC. G. C.										
AH 4-775	T I P Q S L D S W W T S L N F L G G T T V C L G Q N S Q S P T S N H S P T S C P	C.										
AH 4-778	T I P Q S L D S W W T S L N F L G G T T V C L G Q N S Q S P T S N H S P T S C P	C.										
AH 4-2307	T I P Q S L D S W W T S L N F L G G T T V C L G Q N S Q S P T S N H S P T S C P	C.										
AH 4-3445	T I P Q S L D S W W T S L N F L G G T T V C L G Q N S Q S P T S N H S P T S C P	C.										
AH 5. I	T I P Q S L D S W W T S L N F L G G T T V C L G P N S Q S P T S N H S P T S C P	C.										
AH 5-556	T I P Q S L D S W W T S L N F L G G T T V C L G Q N S Q S P T S N H S P T S C P	C.										
AH 5-920	T I P Q S L D S W W T S L N F L G G T T V C L G Q N S Q S P T S N H S P T S C P	C.										
AH 5-1976	T I P Q S L D S W W T S L N F L G G T T V C L G Q N S Q S P T S N H S P T S C P	C.										
AH 6. I	T I P Q S L D S W W T S L N F L G G T T V C L G Q N S Q S P T S N H S P T S C P	C.										
AH 6-217	T I P Q S L D S W W T S L N F L G G T T V C L G Q N S Q S P T S N H S P T S C P	C.										
AH 6-220	T I P Q S L D S W W T S L N F L G G T T V C L G Q N S Q S P T S N H S P T S C P	C.										
AH 7-283	T I P Q S L D S W W T S L N F L G G T T V C L G Q N S Q S P T S N H S P T S C P	C.										

	240	250	260	270	280	290	300	310	320	330	340	350
V01460	ACAATACCGCAGAGTCTAGACTCGTGGTGGACTTCTCTCAATTTTCTAGGGGGAAGTACCGTGTGTCTGGCCAAAATTCGCAGTCCCCAACCTCCAATCACTCACCAACCTTTGTCCCT											
X02496	T I P Q S L D S W W T S L N F L G G T T V C L G Q N S Q S P T S N H S P T S C P											
AH 7-496	T I P Q S L D S W W T S L N F L G G T T V C L G Q N S Q S P T S N H S P T S C P											
AH 7-1010	T I P Q S L D S W W T S L N F L G G T T V C L G Q N S Q S P T S N H S P T S C P											
AH 7-1275	T I P Q S L D S W W T S L N F L G G T T V C L G Q N S Q S P T S N H S P T S C P											
AH 7-1442	T I P Q S L D S W W T S L N F L G G T T V C L G Q N S Q S P T S N H S P T S C P											
AH 8-2186	T I P Q S L D S W W T S L N F L G G T T V C L G Q N S Q S P T S N H S P T S C P											
AH 9-2772	T I P Q S L D S W W T S L N F L G G T T V C L G Q N S Q S P T S N H S P T S C P											
AH 9-296	T I P Q S L D S W W T S L N F L G G T T V C L G Q N S Q S P T S N H S P T S C P											
AH 9-297	T I P Q S L D S W W T S L N F L G G T T V C L G Q N S Q S P T S N H S P T S C P											
AH 9-3163	T I P Q S L D S W W T S L N F L G G T T V C L G Q N S Q S P T S N H S P T S C P											
AH 9-3113	T I P Q S L D S W W T S L N F L G G T T V C L G Q N S Q S P T S N H S P T S C P											
AH 9-3897	T I P Q S L D S W W T S L N F L G G T T V C L G Q N S Q S P T S N H S P T S C P											
AH 9-4070	T I P Q S L D S W W T S L N F L G G T T V C L G Q N S Q S P T S N H S P T S C P											
AH 10-486	T I P Q S L D S W W T S L N F L G G T T V C L G Q N S Q S P T S N H S P T S C P											
AH 10-737	T I P Q S L D S W W T S L N F L G G T T V C L G Q N S Q S P T S N H S P T S C P											
AH 10-885	T I P Q S L D S W W T S L N F L G G T T V C L G Q N S Q S P T S N H S P T S C P											
AH 10-1108	T I P Q S L D S W W T S L N F L G G T T V C L G Q N S Q S P T S N H S P T S C P											
AH 10-1612	T I P Q S L D S W W T S L N F L G G T T V C L G Q N S Q S P T S N H S P T S C P											
AH 10-2085	T I P Q S L D S W W T S L N F L G G T T V C L G Q N S Q S P T S N H S P T S C P											
AH 10-2425	T I P Q S L D S W W T S L N F L G G T T V C L G Q N S Q S P T S N H S P T S C P											
AH 11-965	T I P Q S L D S W W T S L N F L G G T T V C L G Q N S Q S P T S N H S P T S C P											
AH 11-1941	T I P Q S L D S W W T S L N F L G G T T V C L G Q N S Q S P T S N H S P T S C P											
AH 11-2165	T I P Q S L D S W W T S L N F L G G T T V C L G Q N S Q S P T S N H S P T S C P											
AH 11-2514	T I P Q S L D S W W T S L N F L G G T T V C L G Q N S Q S P T S N H S P T S C P											
AH 11-2759	T I P Q S L D S W W T S L N F L G G T T V C L G Q N S Q S P T S N H S P T S C P											
AH 12-116	T I P Q S L D S W W T S L N F L G G T T V C L G Q N S Q S P T S N H S P T S C P											
CH 1-1546	T I P Q S L D S W W T S L N F L G G T T V C L G Q N S Q S P T S N H S P T S C P											
CH 6-2107	T I P Q S L D S W W T S L N F L G G T T V C L G Q N S Q S P T S N H S P T S C P											
CH 7-1117A	T I P Q S L D S W W T S L N F L G G T T V C L G Q N S Q S P T S N H S P T S C P											
CH 9-599	T I P Q S L D S W W T S L N F L G G T T V C L G Q N S Q S P T S N H S P T S C P											
CH 11-884	T I P Q S L D S W W T S L N F L G G T T V C L G Q N S Q S P T S N H S P T S C P											
CH 11-2343A	T I P Q S L D S W W T S L N F L G G T T V C L G Q N S Q S P T S N H S P T S C P											
CH 12-958	T I P Q S L D S W W T S L N F L G G T T V C L G Q N S Q S P T S N H S P T S C P											
COMPL 4-1725	T I P Q S L D S W W T S L N F L G G T T V C L G Q N S Q S P T S N H S P T S C P											
COMPL 11-3028	T I P Q S L D S W W T S L N F L G G T T V C L G Q N S Q S P T S N H S P T S C P											
COMPL 11-3179	T I P Q S L D S W W T S L N F L G G T T V C L G Q N S Q S P T S N H S P T S C P											
X75657	ACAATACCGCAGAGTCTAGACTCGTGGTGGACTTCTCTCAATTTTCTAGGGGGAAGTACCGTGTGTCTGGCCAAAATTCGCAGTCCCCAACCTCCAATCACTCACCAACCTTTGTCCCT											
AH 2-10101	T I P Q S L D S W W T S L N F L G G A P V C L G Q N S Q S P T S N H S P T S C P											

	360	370	380	390	400	410	420	430	440	450	460	470
X02763	CAAATTTCTCCTGGTTATCGCTGGATGTCTCGGGCGTTTTATCATATTCCTCTTCATCCTGCTGCTATGCCCTCATCTTCTTATTGGTTCTTCTGGATTATCAAGGTATGTTGCCCGTT											
AH 1. II	P I C P G Y R W M C L R R F I I F L F I L L L C L I F L L V L L D Y Q G M L P V											
AH 4. I	P I C P G Y R W M C L R R F I I F L F I L L L C L I F L L V L L D Y Q G M L P V											
AH 4-1835	P I C P G Y R W M C L R R F I I F L F I L L L C L I F L L V L L D Y Q G M L P V											
AH 4-2015	P I C P G Y R W M C L R R F I I F L F I L L L C L I F L L V L L D Y Q G M L P V											
AH 5-2878	P I C P G Y R W M C L R R F I I F L F I L L L C L I F L L V L L D Y Q G M L P V											
AH 7-624	P I C P G Y R W M C L R R F I I F L F I L L L C L I F L L V L L D Y Q G M L P V											
AH 7-640	P I C P G Y R W M C L R R F I I F L F I L L L C L I F L L V L L D Y Q G M L P V											
AH 9-454	.C. P T C P G Y R W M C L R R F I I F L F I L L L C L I F L L V L L D Y Q G M L P V											
AH 7-1884	P I C P G Y R W M C L R R F I I F L F I L L L C L I F L L V L L D Y Q G M L P V											
AH 10-2400	P I C P G Y R W M C L R R F I I F L F I L L L C L I F L L V L L D Y Q G M L P V											
AH 10-2596	.A. P I C P G Y R W M C L R R F I I F L F I L L L C L I F L L V L L D Y Q G M L P V											
CH 3-3608	P I C P G Y R W M C L R R F I I F L F I L L L C L I F L L V L L D Y Q G M L P V											
CH 5-2628	P I C P G Y R W M C L R R F I I F L F I L L L C L I F L L V L L D Y Q G M L P V											
CH 6-444A	P I C P G Y R W M C L R R F I I F L F I L L L C L I F L L V L L D Y Q G M L P V											
CH 11-913	P I C P G Y R W M C L R R F I I F L F I L L L C L I F L L V L L D Y Q G M L P V											
CH 12-82	P I C P G Y R W M C L R R F I I F L F I L L L C L I F L L V L L D Y Q G M L P V											
CH 12-2805	.G. P I C P G Y R W M C L R R F I I F L F I L L L C L I F L L V L L D Y Q G M L P V											
COMPL 4-2971	P I C P G Y R W M C L R R F I I F L F I L L L C L I F L L V L L D Y Q G M L P V											
COMPL 9-454	.C. P T C P G Y R W M C L R R F I I F L F I L L L C L I F L L V L L D Y Q G M L P V											
COMPL 9-1553	.C. P I C P G Y R W M C L R R F I I F L F I L L L C L I S L L V L L D Y Q G M L P V											
V01460	CCAACTTCTCCTGGTTATCGCTGGATGTCTCGGGCGTTTTATCATATTCCTCTTCATCCTGCTGCTATGCCCTCATCTTCTTGTGGTTCTTCTGGACTATCAAGGTATGTTGCCCGTT											
X02496	P T C P G Y R W M C L R R F I I F L F I L L L C L I F L L V L L D Y Q G M L P V											
AH 2-2658	P T C P G Y R W M C L R R F I I F L F I L L L C L I F L L V L L D Y Q G M L P V											
AH 3. I	P T C P G Y R W M C L R R F I I F L F I L L L C L I F L L V L L D Y Q G M L P V											
AH 3-577	P T C P G Y R W M C L R R F I I F L F I L L L C L I F L L V L L D Y Q G M L P V											
AH 3-987	P T C P G Y R W M C L R R F I I F L F I L L L C L I F L L V L L D Y Q G M L P V											
AH 3-7755	P T C P G Y R W M C L R R F I I F L F I L L L C L I F L L V L L D Y Q G M L P V											
AH 3-7991	.T. P T C P G Y R W M C L R R F I I F L F I L L L C L I F L L V L L D Y Q G M L P V											
AH 4-775	P T C P G Y R W M C L R R F I I F L F I L L L C L I F L L V L L D Y Q G M L P V											
AH 4-778	P T C P G Y R W M C L R R F I I F L F I L L L C L I F L L V L L D Y Q G M L P V											
AH 4-2307	P T C P G Y R W M C L R R F I I F L F I L L L C L I F L L V L L D Y Q G M L P V											
AH 4-3445	P T C P G Y R W M C L R R F I I F L F I L L L C L I F L L V L L D Y Q G M L P V											
AH 5. I	P T C P G Y R W M C L R R F I I F L F I L L L C L I F L L V L L D Y Q G M L P V											
AH 5-556	P T C P G Y R W M C L R R F I I F L F I L L L C L I F L L V L L D Y Q G M L P V											
AH 5-920	P T C P G Y R W M C L R R F I I F L F I L L L C L I F L L V L L D Y Q G M L P V											
AH 5-1976	P T C P G Y R W M C L R R F I I F L F I L L L C L I F L L V L L D Y Q G M L P V											
AH 6. I	P T C P G Y R W M C L R R F I I F L F I L L L C L I F L L V L L D Y Q G M L P V											
AH 6-217	P T C P G Y R W M C L R R F I I F L F I L L L C L I F L L V L L D Y Q G M L P V											
AH 6-220	P T C P G Y R W M C L R R F I I F L F I L L L C L I F L L V L L D Y Q G M L P V											
AH 7-283	.T. P T C P G Y R W M C L R R F I I F L F I L L L C L I F L L V L L D Y Q G M L P V											

```

          360      370      380      390      400      410      420      430      440      450      460      470
V01460  CCAACTTGTCCCTGGTTATCGCTGGATGTGTCCTGCGGCGTTTTATCATCTTCCCTTCATCCCTGCTGCTATGCCCTCATCTTCTTGTGGTTCTTCTGGACTATCAAGGTATGTTGCCCGTT
X02496  P T C P G Y R W M C L R R F I I F L F I L L L C L I F L L V L L D Y Q G M L P V
AH 7-496  P T C P G Y R W M C L R R F I I F L F I L L L C L I F L L V L L D Y Q G M L P V
AH 7-1010 P T C P G Y R W M C L R R F I I F L F I L L L C L I F L L V L L D Y Q G M L P V
AH 7-1275 P T C P G Y R W M C L R R F I I F L F I L L L C L I F L L V L L D Y Q G M L P V
AH 7-1442 P T C P G Y R W M C L R R F I I F L F I L L L C L I F L L V L L D Y Q G M L P V
AH 8-2186 P T C P G Y R W M C L R R F I I F L F I L L L C L I F L L V L L D Y Q G M L P V
AH 9-2772 P T C P G Y R W M C L R R F I I F L F I L L L C L I F L L V L L D Y Q G M L P V
AH 9-296  P T C P G Y R W M C L R R F I I F L F I L L L C L I F L L V L L D Y Q G M L P V
AH 9-297  P T C P G Y R W M C L R R F I I F L F I L L L C L I F L L V L L D Y Q G M L P V
AH 9-3163 P T C P G Y R W M C L R R F I I F L F I L L L C L I F L L V L L D Y Q G M L P V
AH 9-3113 P T C P G Y R W M C L R R F I I F L F I L L L C L I F L L V L L D Y Q G M L P V
AH 9-3897 P T C P G Y R W M C L R R F I I F L F I L L L C L I F L L V L L D Y Q G M L P V
AH 9-4070 P T C P G Y R W M C L R R F I I F L F I L L L C L I F L L V L L D Y Q G M L P V
AH 10-486 P T C P G Y R W M C L R R F I I F L F I L L L C L I F L L V L L D Y Q G M L P V
AH 10-737 P T C P G Y R W M C L R R F I I F L F I L L L C L I F L L V L L D Y Q G M L P V
AH 10-885 P T C P G Y R W M C L R R F I I F L F I L L L C L I F L L V L L D Y Q G M L P V
AH 10-1108 P T C P G Y R W M C L R R F I I F L F I L L L C L I F L L V L L D Y Q G M L P V
AH 10-1612 P T C P G Y R W M C L R R F I I F L F I L L L C L I F L L V L L D Y Q G M L P V
AH 10-2085 P T C P G Y R W M C L R R F I I F L F I L L L C L I F L L V L L D Y Q G M L P V
AH 10-2425 P T C P G Y R W M C L R R F I I F L F I L L L C L I F L L V L L D Y Q G M L P V
AH 11-965  P T C P G Y R W M C L R R F I I F L F I L L L C L I F L L V L L D Y Q G M L P V
AH 11-1941 P T C P G Y R W M C L R R F I I F L F I L L L C L I F L L V L L D Y Q G M L P V
AH 11-2165 P T C P G Y R W M C L R R F I I F L F I L L L C L I F L L V L L D Y Q G M L P V
AH 11-2514 P T C P G Y R W M C L R R F I I F L F I L L L C L I F L L V L L D Y Q G M L P V
AH 11-2759 P T C P G Y R W M C L R R F I I F L F I L L L C L I F L L V L L D Y Q G M L P V
AH 12-116 P T C P G Y R W M C L R R F I I F L F I L L L C L I F L L V L L D Y Q G M L P V
CH 1-1546 P T C P G Y R W M C L R R F I I F L F I L L L C L I F L L V L L D Y Q G M L P V
CH 6-2107 P T C R G Y R W M C L R H F I I F L F I L L L C L I F L L V L L D Y Q G M L P V
CH 7-1117A P T C P G Y R W M C L R R F I I F L F I L L L C L I F L L V L L D Y Q G M L P V
CH 9-599  P T C P G Y R W M C L R R F I I F L F I L L L C L I F L L V L L D Y Q G M L P V
CH 11-884 P T C P G Y R W M C L R R F I I F L F I L L L C L I F L L V L L D Y Q G M L P V
CH 11-2343 P T C P G Y R W M C L R R F I I F L F I L L L C L I F L L V L L D Y Q G M L P V
CH 12-958 P T C P G Y R W M C L R R F I I F L F I L L L C L I F L L V L L D Y Q G M L P V
COMPL 4-1725 P T C P G Y R W M C L R R F I I F L F I L L L C L I F L L V L L D Y Q G M L P V
COMPL 11-3028 P T C P G Y R W M C L R R F I I F L F I L L L C L I F L L V L L D Y Q G M L P V
COMPL 11-3179 P T C P G Y R W M C L R R F I I F L F I L L L C L I F L L V L L D Y Q G M L P V
X75657  CCAATTTGTCCCTGGCTATCGCTGGATGTGTCCTGCGGCGTTTTATCATCTTCCCTTCATCCCTGCTGCTATGCCCTCATCTTCTTGTGGTTCTTCTGGACTATCAAGGTATGTTGCCCGTT
AH 2-10101 P I C P G Y R W M C L R R F I I F L F I L L L C L I F L L V L L D Y Q G M L P V

```

	480	490	500	510	520	530	540	550	560	570	580	590
X02763	TGTCCTCTAATTCAGGATCAACAACAACAGTACGGGACCATGCAAAACCTGCACGACTCCTGCTCAAGGCAACTCTATGTTTCCCTCATGTGTCTGTACAAAACCTACGGATGAAAT											
AH 1. II	C P L I P G S T T T S T G P C K T C T T P A Q G N S M F P S C C C T K P T D G N											
AH 4. I	C P L I P G S T T T S T G P C K T C T T P A Q G N S M F P S C C C T K P T D G N											
AH 4-1835	C P L I P G S T T T S T G P C K T C T T P A Q G N S M F P S C C C T K P T D G N											
AH 4-2015	C P L I P G S T T T S T G P C K T C T T P A Q G N S M F P S C C C T K P T D G N											
AH 5-2878	C P L I P G S T T T S T G P C K T C T T P A Q G N S M F P S C C C T K P T D G N											
AH 7-624	C P L I P G S T T T S T G P C K T C T T P A Q G N S M F P S C C C T K P T D G N											
AH 7-640	C P L M P G S T T T S T G P C K T C T T P A Q G N S T L P S C C C T K P T D G N											
AH 9-454	C P L I P G S T T T S T G P C K T C T T P A Q G N S M F P S C C C T K P T D G N											
AH 7-1884	C P L I P G S T T T S T G P C K T C T T P A Q G N S M F P S C C C T K P T D G N											
AH 10-2400	C P L I P G S T T T S T G P C K T C T T P A Q G N S M F P S C C C T K P T D G N											
AH 10-2596	C P L I P G S T T T S T G P C K T C T T P A Q G N S M F P S C C C T K P T D G N											
CH 3-3608	C P L I P G S T T T S T G P C K T C T T P A Q G N S M F P S C C C T K P T D G N											
CH 5-2628	C P L I P G S T T T S T G P C K T C T T P A Q G N S M F P S C C C T K P T D G N											
CH 6-444A	C P L I P G S T T T S T G P C K T C T T P A Q G N S M F P S C C C T K P S D G N											
CH 11-913	C P L I P G S T T T S T G P C K T C T T P A Q G N S M F P S C C C T K P T D G N											
CH 12-82	C P L I P G S T T T S T G P C K T C T T P A Q G N S M F P S C C C T K P T D G N											
CH 12-2805	C P L I P G S T T T S T G P C K T C T T P A Q G N S M F P S C C C T K P S D G N											
COMPL 4-2971	C P L I P G S T T T S T G P C K T C T T P A Q G N S M F P S C C C T K P T D G N											
COMPL 9-454	C P L I P G S T T T S T G P C K T C T T P A Q G N S M F P S C C C T K P T D G N											
COMPL 9-1553	C P L I P G S T T T S T G P C K T C T T P A Q G N S M F P S C C C T K P S D G N											
V01460	TGTCCTCTAATTCAGGATCCCAACAACAGCAGGACCATGCCGACCTGCATGACTACTGCTCAAGGCAACTCTATGTATCCCTCTGTGTCTGTACAAAACCTTCGGACGGAAT											
X02496	C P L I P G S S T T S T G P C R T C M T T A Q G T S M Y P S C C C T K P S D G N											
AH 2-2658	C P L I P G S S T T S T G P C R T C T T P A Q G T S M Y P S C C C T K P S D G N											
AH 3. I	C P L I P G S S T T S V G P C R T C T T T V Q G T S M Y P S C C C T K P S D G N											
AH 3-577	C P L I P G S S T T S V G P C R T C T T T V Q G T S M Y P S C C C T K P S D G N											
AH 3-987	C P L I P G S S T T S A G P C R T C T T T A Q G T S I Y P S C C C T K P S D G N											
AH 3-7755	C P L I P G S S T T S V G P C R T C T T T V Q G T S M Y P S C C C T K P S D G N											
AH 3-7991	C P L I P G S S T T S T G P C R T C T T P A Q G T S M Y P S C C C T K P S D G N											
AH 4-775	C P L I P G S S T T S A G P C R T C T T T A Q G T S I Y P S C C C T K P S D G N											
AH 4-778	C P L I P G S S T T S A G P C R T C T T T A Q G T S M Y P S C C C T K P S D G N											
AH 4-2307	C P L I P G S S T T S M G P C R T C T T T V Q G T S M Y P S C C C T K P S D G N											
AH 4-3445	C P L I P G S S T T S T G P C R T C T T P A Q G T S M Y P S C C C T K P S D G N											
AH 5. I	C P L I P G S S T T S V G P C R T C T T T V Q G T S M Y P S C C C T K P S D G N											
AH 5-556	C P L I P G S S T T S A G P C R T C T T T A Q G T S M Y P S C C C T K P S D G N											
AH 5-920	C P L I P G S S T T S V G P C R T C T T T V Q G T S M Y P S C C C T K P S D G N											
AH 5-1976	C P L I P G S S T T S V G P C R T C T T T V Q G T S M Y P S C C C T K P S D G N											
AH 6. I	C P L I P G S S T T S V G P C R T C T T T V Q G T S M Y P S C C C T K P S D G N											
AH 6-217	C P L I P G S S T T S A G P C R T C T T T A Q G T S M Y P S C C C T K P S D G N											
AH 6-220	C P L I P G S S T T S V G P C R T C T T T V Q G T S M Y P Y C C C T K P S D G N											
AH 7-283	C P L I P G S S T T S T G P C R T C T T T A Q G T S M Y P S C C C T K P S D G N											

	480	490	500	510	520	530	540	550	560	570	580	590																												
V01460	TGTCCTCTAATCCAGGATCCCTCAACAACCAGCACGGGACCATGCCGGACCTGCATGACTACTGCTCAAGGAACCTCTATGTATCCCTCCTGTTGCTGTACCAAAACCTTCGGACGGAAAT																																							
X02496	C	P	L	I	P	G	S	S	T	T	S	T	G	P	C	R	T	C	M	T	T	A	Q	G	T	S	M	Y	P	S	C	C	C	T	K	P	S	D	G	N
AH 7-496	C	P	L	I	P	G	S	S	T	T	S	T	G	P	C	R	T	C	T	T	P	A	Q	G	T	S	M	Y	P	S	C	C	C	T	K	P	S	D	G	N
AH 7-1010	C	P	L	I	P	G	S	S	T	T	S	T	G	P	C	R	T	C	T	T	P	A	Q	G	T	S	M	Y	P	S	C	C	C	T	K	P	S	D	G	N
AH 7-1275	C	P	L	I	P	G	S	S	T	T	S	T	G	P	C	R	T	C	T	T	T	A	Q	G	T	S	M	Y	P	S	C	C	C	T	K	P	S	D	G	N
AH 7-1442	C	P	L	I	P	G	S	S	T	T	S	T	G	P	C	R	T	C	T	T	P	A	Q	G	T	S	M	Y	P	S	C	C	C	T	K	P	S	D	G	N
AH 8-2186	C	P	L	I	P	G	S	S	T	T	S	T	G	P	C	R	T	C	T	T	T	A	Q	G	T	S	M	Y	P	S	C	C	C	T	K	P	S	D	G	N
AH 9-2772	C	P	L	I	P	G	S	S	T	T	S	T	G	P	C	R	T	C	T	T	P	A	Q	G	T	S	M	Y	P	S	C	C	C	T	K	P	S	D	G	N
AH 9-296	C	P	L	I	P	G	S	S	T	T	S	T	G	P	C	R	T	C	T	T	T	A	Q	G	T	S	M	Y	P	S	C	C	C	T	K	P	S	D	G	N
AH 9-297	C	P	L	I	P	G	S	S	T	T	S	T	G	P	C	R	T	C	T	T	T	A	Q	G	T	S	M	Y	P	S	C	C	C	T	K	P	S	D	G	N
AH 9-3163	C	P	L	I	P	G	S	S	T	T	S	T	G	P	C	R	T	C	T	T	T	V	Q	G	T	S	M	Y	P	S	C	C	C	T	K	P	S	D	G	N
AH 9-3113	C	P	L	I	P	G	S	S	T	T	S	T	G	P	C	R	T	C	T	T	T	A	Q	G	T	S	M	Y	P	S	C	C	C	T	K	P	S	D	G	N
AH 9-3897	C	P	L	I	P	G	S	S	T	T	S	T	G	P	C	R	T	C	T	T	T	V	Q	G	T	S	M	Y	P	S	C	C	C	T	K	P	S	D	G	N
AH 9-4070	C	P	L	I	P	G	S	S	T	T	S	T	G	P	C	R	T	C	T	T	T	A	Q	G	T	S	M	Y	P	S	C	C	C	T	K	P	S	D	G	N
AH 10-486	C	P	L	I	P	G	S	S	T	T	S	T	G	P	C	R	T	C	T	T	T	A	Q	G	T	S	M	Y	P	S	C	C	C	T	K	P	S	D	G	N
AH 10-737	C	P	L	I	P	G	S	S	T	T	S	T	G	P	C	R	T	C	T	T	T	A	Q	G	T	S	M	Y	P	S	C	C	C	T	K	P	S	D	G	N
AH 10-885	C	P	L	I	P	G	S	S	T	T	S	T	G	P	C	R	T	C	T	T	T	V	Q	G	T	S	M	Y	P	S	C	C	C	T	K	P	S	D	G	N
AH 10-1108	C	P	L	I	P	G	S	S	T	T	S	T	G	P	C	R	T	C	T	T	T	A	Q	G	T	S	M	Y	P	S	C	C	C	T	K	P	S	D	G	N
AH 10-1612	C	P	L	I	P	G	S	S	T	T	S	T	G	P	C	R	T	C	T	T	T	V	Q	G	T	S	M	Y	P	S	C	C	C	T	K	P	S	D	G	N
AH 10-2085	C	P	L	I	P	G	S	S	T	T	S	T	G	P	C	R	T	C	T	T	T	V	Q	G	T	S	M	Y	P	S	C	C	C	T	K	P	S	D	G	N
AH 10-2425	C	P	L	I	P	G	S	S	T	T	S	T	G	P	C	R	T	C	T	T	T	A	Q	G	T	S	M	Y	P	S	C	C	C	T	K	P	S	D	G	N
AH 11-965	C	P	L	I	P	G	S	S	T	T	S	T	G	P	C	R	T	C	T	T	T	A	Q	G	T	S	M	Y	P	S	C	C	C	T	K	P	S	D	G	N
AH 11-1941	C	P	L	I	P	G	S	S	T	T	S	T	G	P	C	R	T	C	T	T	P	A	Q	G	T	S	M	Y	P	S	C	C	C	T	K	P	S	D	G	N
AH 11-2165	C	P	L	I	P	G	S	S	T	T	S	T	G	P	C	R	T	C	T	T	T	A	Q	G	T	S	M	Y	P	S	C	C	C	T	K	P	S	D	G	N
AH 11-2514	C	P	L	I	P	G	S	S	T	T	S	T	G	P	C	R	T	C	T	T	T	V	Q	G	T	S	M	Y	P	S	C	C	C	T	K	P	S	D	G	N
AH 11-2759	C	P	L	I	P	G	S	S	T	T	S	T	G	P	C	R	T	C	T	T	T	V	Q	G	T	S	M	Y	P	S	C	C	C	T	K	P	S	D	G	N
AH 12-116	C	P	L	I	P	G	S	S	T	T	S	T	G	P	C	R	T	C	T	T	T	V	Q	G	T	S	M	Y	P	S	C	C	C	T	K	P	S	D	G	N
CH 1-1546	C	P	L	I	P	G	S	S	T	T	S	T	G	P	C	R	T	C	T	T	S	A	Q	G	T	S	M	K	P	S	C	C	C	T	K	P	S	D	G	N
CH 6-2107	C	P	L	I	P	G	S	S	T	T	S	T	G	P	C	R	T	C	T	T	P	A	Q	G	T	S	M	Y	P	S	C	C	C	T	K	P	S	D	G	N
CH 7-1117A	C	P	L	I	P	G	S	S	T	T	S	T	G	P	C	R	T	C	T	T	P	A	Q	G	T	S	M	Y	P	S	C	C	C	T	K	P	S	D	G	N
CH 9-599	C	P	L	I	P	G	S	S	T	T	S	T	G	P	C	R	T	C	T	T	P	A	Q	G	T	S	M	Y	P	S	C	C	C	T	K	P	S	D	G	N
CH 11-884	C	P	L	I	P	G	S	S	T	T	S	T	G	P	C	R	T	C	T	T	P	A	Q	G	T	S	M	Y	P	S	C	C	C	T	K	P	S	D	G	N
CH 11-2343A	C	P	L	I	P	G	S	S	T	T	S	T	G	P	C	R	T	C	T	T	P	A	Q	G	T	S	M	Y	P	S	C	C	C	T	K	P	S	D	G	N
CH 12-958	C	P	L	I	P	G	S	S	T	T	S	T	G	P	C	R	T	C	T	T	T	A	Q	G	T	S	M	Y	P	S	C	C	C	T	K	P	S	D	G	N
COMPL 4-1725	C	P	L	I	P	G	S	S	T	T	S	T	G	P	C	R	T	C	T	T	T	A	Q	G	T	S	M	Y	P	S	C	C	C	T	K	P	S	D	G	N
COMPL 11-3028	C	P	L	I	P	G	S	S	T	T	S	T	G	P	C	R	T	C	T	T	T	A	Q	G	T	S	M	Y	P	S	C	C	C	T	K	P	S	D	G	N
COMPL 11-3179	C	P	L	I	P	G	S	S	T	T	S	T	G	P	C	R	T	C	T	T	T	A	Q	G	T	S	M	Y	P	S	C	C	C	T	K	P	S	D	G	N
X75657	TGTCCTCTAATCCAGGATCATCAACCACAGTACGGGACCCCTGCCGAACCTGCACGACTCTTGCTCAAGGAACCTCTATGTTCCCTCATGTTGCTGTTCACAAACCTTCGGACGGAAAT																																							
AH 2-10101	C	P	L	I	P	G	S	S	T	T	S	T	G	P	C	R	T	C	T	T	L	A	Q	G	T	S	M	F	P	S	C	C	C	S	K	P	S	D	G	N
	C	P	L	I	P	G	S	S	T	T	S	T	G	P	C	R	T	C	T	T	L	A	Q	G	T	S	M	F	P	S	C	C	C	S	K	P	S	D	G	N

	600	610	620	630	640	650	660	670	680	690	700	710																												
X02763	TGCACCTGATTCCCATCCCATCGTCTGGGCTTTCGCAAAATACCTATGGGAGTGGGCCTCAGTCCGTTTCTCTGGCTCAGTTTACTAGTGCCATTGTTCAGTGGTTTCGTAGGGCTT																																							
AH 1. II	C	T	C	I	P	I	P	S	S	W	A	F	A	K	Y	L	W	E	W	A	S	V	R	F	S	W	L	S	L	L	V	P	F	V	Q	W	F	V	G	L
AH 4. I	C	T	C	I	P	I	P	S	S	W	A	F	A	K	Y	L	W	E	W	A	S	V	R	F	S	W	L	S	L	L	V	P	F	V	Q	W	F	V	G	L
AH 4-1835	C	T	C	I	P	I	P	S	S	W	A	F	A	K	Y	L	W	E	W	A	S	V	R	F	S	W	L	S	L	L	V	P	F	V	Q	W	F	V	G	L
AH 4-2015	C	T	C	I	P	I	P	S	S	W	A	F	A	K	Y	L	W	E	W	A	S	A	R	F	S	W	L	S	L	L	V	P	F	V	Q	W	F	V	G	L
AH 5-2878	C	T	C	I	P	I	P	S	S	W	A	F	A	K	Y	L	W	E	W	A	S	V	R	F	S	W	L	S	L	L	V	P	F	V	Q	W	F	V	G	L
AH 7-624	C	T	C	I	P	I	P	S	S	W	A	F	A	K	Y	L	W	E	W	A	S	V	R	F	S	W	L	S	L	L	V	P	F	V	Q	W	F	V	G	L
AH 7-640	C	T	C	I	P	I	P	S	S	W	A	F	A	K	Y	L	W	E	W	A	S	V	R	F	S	W	L	S	L	L	V	P	F	V	Q	W	F	V	G	L
AH 9-454	C	T	C	I	P	I	P	S	S	W	A	F	A	K	Y	L	W	E	W	A	S	V	R	F	S	W	L	S	L	L	V	P	F	V	Q	W	F	V	G	L
AH 7-1884	C	T	C	I	P	I	P	S	S	W	A	F	A	K	Y	L	W	E	W	A	S	V	R	F	S	W	L	S	L	L	V	P	F	V	Q	W	F	V	G	L
AH 10-2400	C	T	C	I	P	I	P	S	S	W	A	F	A	K	Y	L	W	E	W	A	S	V	R	F	S	W	L	S	L	L	V	P	F	V	Q	W	F	V	G	L
AH 10-2596	C	T	C	I	P	I	P	S	S	W	A	F	A	K	Y	L	W	E	W	A	S	V	R	F	S	W	L	S	L	L	V	P	F	V	Q	W	F	V	G	L
CH 3-3608	C	T	C	I	P	I	P	S	S	W	A	F	A	K	Y	L	W	E	W	A	S	V	R	F	S	W	L	S	L	L	V	P	F	V	Q	W	F	V	G	L
CH 5-2628	C	T	C	I	P	I	P	S	S	W	A	F	A	K	Y	L	W	E	W	A	S	V	R	F	S	W	L	S	L	L	V	P	F	V	Q	W	F	V	G	L
CH 6-444A	C	T	C	I	P	I	P	S	S	W	A	F	A	K	Y	L	W	E	W	A	S	V	R	F	S	W	L	S	L	L	V	P	F	V	Q	W	F	V	G	L
CH 11-913	C	T	C	I	P	I	P	S	S	W	A	F	A	K	Y	L	W	E	W	A	S	V	R	F	S	W	L	S	L	L	V	P	F	V	Q	W	F	V	G	L
CH 12-82	C	T	C	I	P	I	P	S	S	W	A	F	A	K	Y	L	W	E	W	A	S	V	R	F	S	W	L	S	L	L	V	P	F	V	Q	W	F	V	G	L
CH 12-2805	C	T	C	I	P	I	P	S	S	W	A	F	A	K	Y	L	W	E	W	A	S	A	R	F	S	W	L	S	L	L	V	P	F	V	Q	W	F	V	G	L
COMPL 4-2971	C	T	C	I	P	I	P	S	S	W	A	F	A	K	Y	L	W	E	W	A	S	V	R	F	S	W	L	S	L	L	V	P	F	V	Q	W	F	V	G	L
COMPL 9-454	C	T	C	I	P	I	P	S	S	W	A	F	A	K	Y	L	W	E	W	A	S	V	R	F	S	W	L	S	L	L	V	P	F	V	Q	W	F	L	G	L
COMPL 9-1553	C	T	C	I	P	I	P	S	S	W	A	F	A	K	Y	L	W	E	W	A	S	V	R	F	S	W	L	S	L	L	V	P	F	V	Q	W	F	V	G	L
V01460	TGCACCTGATTCCCATCCCATCATCTGGGCTTTCGGAATAATCCATATGGGAGTGGGCCTCAGCCCGTTTCTCTGGCTCAGTTTACTAGTGCCATTGTTCAGTGGTTTCGTAGGGCTT																																							
X02496	C	T	C	I	P	I	P	S	S	W	A	F	G	K	F	L	W	E	W	A	S	A	R	F	S	W	L	S	L	L	V	P	F	V	Q	W	F	V	G	L
AH 2-2658	C	T	C	I	P	I	P	S	S	W	A	F	G	K	F	L	W	E	W	A	S	A	R	F	S	W	L	S	L	L	V	P	F	V	Q	W	F	V	G	L
AH 3. I	C	T	C	I	P	I	P	S	S	W	A	F	G	K	F	L	W	E	W	A	S	A	R	F	S	W	L	S	L	L	V	P	F	V	Q	W	F	V	G	L
AH 3-577	C	T	C	I	P	I	P	S	S	W	A	F	G	K	F	L	W	E	W	A	S	A	R	F	S	W	L	S	L	L	V	P	F	V	Q	W	F	V	G	L
AH 3-987	C	T	C	I	P	I	P	S	S	W	A	F	G	K	F	L	W	E	W	A	S	A	R	F	S	W	L	S	L	L	V	P	F	V	Q	W	F	V	G	L
AH 3-7755	C	T	C	I	P	I	P	S	S	W	A	F	G	K	F	L	W	E	W	A	S	A	R	F	S	W	L	S	L	L	V	P	F	V	Q	W	F	V	G	L
AH 3-7991	C	T	C	I	P	I	P	S	S	W	A	F	G	K	F	L	W	E	W	A	S	A	R	F	S	W	L	S	L	L	V	P	F	V	Q	W	F	V	G	L
AH 4-775	C	T	C	I	P	I	P	S	S	W	A	F	G	K	F	L	W	E	W	A	S	A	R	F	S	W	L	S	L	L	V	P	F	V	Q	W	F	V	G	L
AH 4-778	C	T	C	I	P	I	P	S	S	W	A	F	G	K	F	L	W	E	W	A	S	A	R	F	S	W	L	S	L	L	V	P	F	V	Q	W	F	V	G	L
AH 4-2307	C	T	C	I	P	I	P	S	S	W	A	F	G	K	F	L	W	E	W	A	S	A	R	F	S	W	L	S	L	L	V	P	F	V	Q	W	F	V	G	L
AH 4-3445	C	T	C	I	P	I	P	S	S	W	A	F	G	K	F	L	W	E	W	A	S	A	R	F	S	W	L	S	L	L	V	P	F	V	Q	W	F	V	G	L
AH 5. I	C	T	C	I	P	I	P	S	S	W	A	F	G	K	F	L	W	E	W	A	S	A	R	F	S	W	L	S	L	L	V	P	F	V	Q	W	F	V	G	L
AH 5-556	C	T	C	I	P	I	P	S	S	W	A	F	G	K	F	L	W	E	W	A	S	A	R	F	S	W	L	S	L	L	V	P	F	V	Q	W	F	V	G	L
AH 5-920	C	T	C	I	P	I	P	S	S	W	A	F	G	K	F	L	W	E	W	A	S	A	R	F	S	W	L	S	L	L	V	P	F	V	Q	W	F	V	G	L
AH 5-1976	C	T	C	I	P	I	P	S	S	W	A	F	G	K	F	L	W	E	W	A	S	A	R	F	S	W	L	S	L	L	V	P	F	V	Q	W	F	V	G	L
AH 6. I	C	T	C	I	P	I	P	S	S	W	A	F	G	K	F	L	W	E	W	A	S	A	R	F	S	W	L	S	L	L	V	P	F	V	Q	W	F	V	G	L
AH 6-217	C	T	C	I	P	I	P	S	S	W	A	F	G	K	F	L	W	E	W	A	S	A	R	F	S	W	L	S	L	L	V	P	F	V	Q	W	F	V	G	L
AH 6-220	C	T	C	I	P	I	P	S	S	W	A	F	G	K	F	L	W	E	W	A	S	A	R	F	S	W	L	S	L	L	V	P	F	V	Q	W	F	V	G	L
AH 7-283	C	T	C	I	P	I	P	S	S	W	A	F	G	K	F	L	W	E	W	A	S	A	R	F	S	W	L	S	L	L	V	P	F	V	Q	W	F	V	G	L

	600	610	620	630	640	650	660	670	680	690	700	710
V01460	TGCACCTGTATTCCCATCCCATCATCTGGGCTTTTCGGAAAATTCCTATGGGAGTGGGCCTCAGCCCGTTTCCTGGCTCAGTTTACTAGTGCCATTGTTTCAGTGGTTTCGATGGGCTT											
X02496	C T C I P I P S S W A F G K F L W E W A S A R F S W L S L L V P F V Q W F V G L											
AH 7-496	C T C I P I P S S W A F G K F L W E W A S A R F S W L S L L V P F V Q W F V G L											
AH 7-1010	C T C I P I P S S W A F G K F L W E W A S A R F S W L S L L V P F V Q W F V G L											
AH 7-1275	C T C I P I P S S W A F G K F L W E W A S A R F S W L S L L V P F V Q W F V G L											
AH 7-1442	C T C I P I P S S W A F G K F L W E W A S A R F S W L S L L V P F V Q W F V G L											
AH 8-2186	C T C I P I P S S W A F G K F L W E W A S A R F S W L S L L V P F V Q W F V G L											
AH 9-2772	C T C I P I P S S W A F G K F L W E W A S A R F S W L S L L V P F V Q W F V G L											
AH 9-296	C T C I P I P S S W A F G K F L W E W A S A R F S W L S L L V P F V Q W F V G L											
AH 9-297	C T C I P I P S S W A F G K F L W E W A S A R F S W L S L L V P F V Q W F V G L											
AH 9-3163	C T C I P I P S S W A F G K F L W E W A S A R F S W L S L L V P F V Q W F V G L											
AH 9-3113	C T C I P I P S S W A F G K F L W E W A S A R F S W L S L L V P F V Q W F V G L											
AH 9-3897	C T C I P I P S S W A F G K F L W E W A S A R F S W L S L L V P F V Q W F V G L											
AH 9-4070	C T C I P I P S S W A F G K F L W E W A S A R F S W L S L L V P F V Q W F V G L											
AH 10-486	C T C I P I P S S W A F G K F L W E W A S A R F S W L S L L V P F V Q W F V G L											
AH 10-737	C T C I P I P S S W A F G K F L W E W A S A R F S W L S L L V P F V Q W F V G L											
AH 10-885	C T C I P I P S S W A F G K F L W E W A S A R F S W L S L L V P F V Q W F V G L											
AH 10-1108	C T C I P I P S S W A F G K F L W E W A S A R F S W L S L L V P F V Q W F V G L											
AH 10-1612	C T C I P I P S S W A F G K F L W E W A S A R F S W L S L L V P F V Q W F V G L											
AH 10-2085	C T C I P I P S S W A F G K F L W E W A S A R F S W L S L L V P F V Q W F V G L											
AH 10-2425	C T C I P I P S S W A F G K F L W E W A S A R F S W L S L L V P F V Q W F V G L											
AH 11-965	C T C I P I P S S W A F G K F L W E W A S A R F S W L S L L V P F V Q W F V G L											
AH 11-1941	C T C I P I P S S W A F G K F L W E W A S A R F S W L S L L V P F V Q W F V G L											
AH 11-2165	C T C I P I P S S W A F G K F L W E W A S A R F S W L S L L V P F V Q W F V G L											
AH 11-2514	C T C I P I P S S W A F G K F L W E W A S A R F S W L S L L V P F V Q W F V G L											
AH 11-2759	C T C I P I P S S W A F G K F L W E W A S A R F S W L S L L V P F V Q W F V G L											
AH 12-116	C T C I P I P S S W A F G K F L W E W A S A R F S W L S L L V P F V Q W F V G L											
CH 1-1546	C T C I P I P S S W A F G K F L W E W G S A R F S W L S L L V P F V Q W F V G L											
CH 6-2107	C T C I P I P S S W A F G K F L W E W A S A R F S W L S L L V P F V Q W F V G L											
CH 7-1117A	C T C I P I P S S W A F G K F L W E W A S A R F S W L S L L V P F V Q W F V G L											
CH 9-599	C T C I P I P S S W A F G K F L W E W A S A R F S W L S L L V P F V Q W F V G L											
CH 11-884	C T C I P I P S S W A F G K F L W E W A S A R F S W L S L L V P F V Q W F V G L											
CH 11-2343A	C T C I P I P S S W A F G K F L W E W A S A R F S W L S L L V P F V Q W F V G L											
CH 12-958	C T C I P I P S S W A F G K F L W E W A S A R F S W L S L L V P F V Q W F V G L											
COMPL 4-1725	C T C I P I P S S W A F G K F L W E W A S A R F S W L S L L V P F V Q W F V G L											
COMPL 11-3028	C T C I P I P S S W A F G K F L W E W A S A R F S W L S L L V P F V Q W F V G L											
COMPL 11-3179	C T C I P I P S S W A F G K F L W E W A S A R F S W L S L L V P F V Q W F V G L											
X75657	TGCACCTGTATTCCCATCCCATCATCTGGGCTTTTCGGAAAATTCCTATGGGAGTGGGCCTCAGCCCGTTTCCTGGCTCAGTTTACTAGTGCCATTGTTTCAGTGGTTTCGATGGGCTT											
AH 2-10101	C T C I P I P S S W A F G K F L W E W A S A R F S W L S L L V P F V Q W F A G L											

	720	730	740	750	760	770	780	790
X02763	TCCCCACTGTTGGCTTTCAGCTATATGGATGATGGTATTGGGGCCAAGTCTGTACAGCATCGTGAGTCCCTTTATACC							
AH 1.II	S P T V W L S A I W M M W Y W G P S L Y S I V S P F I							
AH 4.I	S P T V W L S A I W M M W Y W G P S L Y S I V S P F I							
AH 4-1835	S P T V W L S A I W M M W Y W G P S L Y S I V S P F I							
AH 4-2015	S P T V W L S A I W M M W Y W G P S L Y S I V S P F I							
AH 5-2878	S P T V W L S A I W M M W Y W G P S L Y S I V S P F I							
AH 7-624	S P T V W L S A I W M M W Y W G P S L Y S I V S P F I							
AH 7-640	S P T V W L S A I W M M W Y W G P S L Y S I V R P F I							
AH 9-454	S P T V W L S A I W M M W Y W G P S L Y S I V S P F I							
AH 7-1884	S P T V W L S V I W M M W Y W G P S L Y S I L S P F L							
AH 10-2400	S P T V W L S A I W M M W Y W G P S L Y S I V S P F I							
AH 10-2596	S P T V W L S A I W M M W Y W G P S L Y S I V S P F I							
CH 3-3608	S P T V W L S A I W M M W Y W G P S L Y S I V S P F I							
CH 5-2628	S P T V W L S A I W M M W Y W G P S L Y S I V S P F I							
CH 6-444A	S P T V W L S A I W M M W Y W G P S L Y S I V S P F I							
CH 11-913	S P T V W L S A I W M M W Y W G P S L Y S I V S P F I							
CH 12-82	S P T V W L S A I W M M W Y W G P S L Y S I V S P F I							
CH 12-2805	S P T V W L S A I W M M W Y W G P S L Y S I V S P F I							
COMPL 4-2971	S P T V W L S A I W M M W Y W G P S L Y S I V S P F I							
COMPL 9-454	S P T V W L S A I W M M W Y W G P S L Y S I V S P F I							
COMPL 9-1553	S P T V W L S A I W M M W Y W G P S L Y S I V S P F I							
V01460	TCCCCACTGTTGGCTTTCAGTTATATGGATGATGGTATTGGGGCCAAGTCTGTACAGCATCTTGAGTCCCTTTTACC							
X02496	S P T V W L S V I W M M W Y W G P S L Y S I L S P F L							
AH 2-2658	S P T V W L S V I W M M W Y W G P S L Y S I L S P F L							
AH 3.I	S P T V W L S V I W M M W Y W G P S L Y S I L S P F L							
AH 3-577	S P T V W L S V I W M M W Y W G P S L Y S I L S P F L							
AH 3-987	S P T V W L S V I W M M W Y W G P S L Y S I L S P F L							
AH 3-7755	S P T V W L S V I W M M W Y W G P S L Y S I L S P F L							
AH 3-7991	S P T V W L S V I W T M W Y W G P N L Y N I L S P F L							
AH 4-775	S P T V W L S V I W M M W Y W G P S L Y S I L S P F L							
AH 4-778	S P T V W L S V I W M M W Y W G P S L Y S I L S P F L							
AH 4-2307	S P T V W L S V I W M M W Y W G P S L Y S I L S P F L							
AH 4-3445	S P T V W L S V I W M M W Y W G P S L Y S I L S P F L							
AH 5.I	S P T V W L S V I W M M W Y W G P S L Y S I L S P F L							
AH 5-556	S P T V W L S V I W M M W Y W G P S L Y S I L S P F L							
AH 5-920	S P T V W L S V I W M M W Y W G P S L Y S I L S P F L							
AH 5-1976	S P T V W L S V I W M M W Y W G P S L Y S I L S P F L							
AH 6.I	S P T V W L S V I W M M W Y W G P S L Y S I L S P F L							
AH 6-217	S P T V W L S V I W M M W Y W G P S L Y S I L S P F L							
AH 6-220	S P T V W L S V I W M M W Y W G P S L Y S I L S P F L							
AH 7-283	S P T V W L S V I W M M W Y W G P S L Y S I L S P F L							

	720	730	740	750	760	770	780	790
V01460	TCCCCACTGTTGGCTTTCAGTTATATGGATGATGGTATTGGGGCCCAAGTCTGTACAGCATCTTGAGTCCCTTTTACC							
X02496	S P T V W L S V I W M M W Y W G P S L Y S I L S P F L							
AH 7-496	S P T V W L S V I W M M W Y W G P S L Y S I L S P F L							
AH 7-1010	S P T V W L S V I W M M W Y W G P S L Y S I L S P F L							
AH 7-1275	S P T V W L S V I W M M W Y W G P S L Y S I L S P F L							
AH 7-1442	S P T V W L S V I W M M W Y W G P S L Y S I L S P F L							
AH 8-2186	S P T V W L S V I W M M W Y W G P S L Y S I L S P F L							
AH 9-2772	S P T V W L S V I W M M W Y W G P S L Y S I L S P F L							
AH 9-296	S P T V W L S V I W M M W Y W G P S L Y S I L S P F L							
AH 9-297	S P T V W L S V I W M M W Y W G P S L Y S I L S P F L							
AH 9-3163	S P T V W L S V I W M M W Y W G P S L Y S I L S P F L							
AH 9-3113	S P T V W L S V I W M M W Y W G P S L Y S I L S P F L							
AH 9-3897	S P T V W L S V I W M M W Y W G P S L Y S I L S P F L							
AH 9-4070	S P T V W L S V I W M M W Y W G P S L Y S I L S P F L							
AH 10-486	S P T V W L S V I W M M W Y W G P S L Y S I L S P F L							
AH 10-737	S P T V W L S V I W M M W Y W G P S L Y S I L S P F L							
AH 10-885	S P T V W L S V I W M M W Y W G P S L Y S I L S P F L							
AH 10-1108	S P T V W L S V I W M M W Y W G P S L Y S I L S P F L							
AH 10-1612	S P T V W L S V I W M M W Y W G P S L Y S I L S P F L							
AH 10-2085	S P T V W L S V I W M M W Y W G P S L Y S I L S P F L							
AH 10-2425	S P T V W L S V I W M M W Y W G P S L Y S I L S P F L							
AH 11-965	S P T V W L S V I W M M W Y W G P S L Y S I L S P F L							
AH 11-1941	S P T V W L S V I W M M W Y W G P S L Y S I L S P F L							
AH 11-2165	S P T V W L S V I W M M W Y W G P S L Y S I L S P F L							
AH 11-2514	S P T V W L S V I W M M W Y W G P S L Y S I L S P F L							
AH 11-2759	S P T V W L S V I W M M W Y W G P S L Y S I L S P F L							
AH 12-116	S P T V W L S V I W M M W Y W G P S L Y S I L S P F L							
CH 1-1546	S P T V W L S V I W M M W Y W G P N L Y S I L S P F L							
CH 6-2107	S P T V W L S V I W M M W Y W G P S L Y S I L S P F L							
CH 7-1117A	S P T V W L S V I W M M W Y W G P S L Y S I L S P F L							
CH 9-599	S P T V W L S V I W M M W Y W G P S L Y S I L S P F L							
CH 11-884	S P T V W L S V I W M M W Y W G P S L Y S I L S P F L							
CH 11-2343A	S P T V W L S V I W M M W Y W G P S L Y S I L S P F L							
CH 12-958	S P T V W L S V I W M M W Y W G P S L Y S I L S P F L							
COMPL 4-1725	S P T V W L S V I W M M W Y W G P S L Y S I L S P F L							
COMPL 11-3028	S P T V W L S V I W M M W Y W G P S L Y S I L S P F L							
COMPL 11-3179	S P T V W L S V I W M M W Y W G P S L Y S I L S P F L							
X75657	TCCCCACTGTCGGCTTTCAGTTATATGGATGATGGTATTGGGGCCCAAGTCTGTACAACATCTTGAGTCCCTTTTACC							
AH 2-10101	S P T V W L S V I W M M W Y W G P S L Y N I L S P F I							

Appendix Part C. Sequences of gene X obtained during the study.

	1510	1520	1530	1540	1550	1560	1570	1580	1590	1600	1610	1620
X02763	GTTCCAGCCGACCACGGGGCGCACCTCTCTTTACGCGGTCTCCCGTCTGTGCCTTCTCATCTGCCGGTCCGGTGTGCACCTTCGCTTCACCTCTGCACGTTGCATGGAGACCACCGTGAAC											
AH 4-2015	V P A D H G A H L S L R G L P V C A F S S A G P C A L R F T S A R C M E T T V N											
AH 7-640	V P A D H G A H L S L R G L P V C A F S S A G P C A L R F T S A R C M E T T V N .C.											
AH 8-1993	V P A D H G A H L S L R G L P V C A F S S A G P C A L R F T S A R C M E T T V N											
CH 5-2628	V P A D H G A H L S L R G L P V C A F S S A G P C A L R F T S A R C M E T T V N											
CH 12-2805	V P A D H G A H L S L R G L P V C A F S S A G P C A L R F T S A R C M E T T V N											
COMPL 4-2971	V P A D H G A H L S L R G L P V C A F S S A G P C A L R F T S A R C M E T T V N											
V01460REV	GTTCCGACCCGACCACGGGGCGCACCTCTCTTTACGCGGTCTCCCGTCTGTGCCTTCTCATCTGCCGGTCCGGTGTGCACCTTCGCTTCACCTCTGCACGTTGCATGGAGACCACCGTGAAC											
X02496	V P T D H G A H L S L R G L P V C A F S S A G P C A L R F T S A R R M E T T V N .A											
AH 2-2658	V S T D H G A H L S L R G L P V C A F S S A G P C A L R F T S A R R M E T T V K											
AH 3-7755	V P T D H G A H L S L R G L P V C A F S S A G P C A L R F T S A R R M E T T V N											
AH 3-7991	V P T D H G A H L S L R G L P V C A F S S A G P C A L R F T S A R R M E T T V N .C. .A.											
AH 3-987	V P T D H G A H L S L R G L P V C A F S S A G P C A L R F T S A R R M E T T V N											
AH 5-556	V P T D H G A H L S L R G L P V C A F S S A G P C A L R F T S A R R M E T T V N											
AH 5-920	V P T D H G A H L S L R G L P V C A F S S A G P C A L R F T S A R R M E T T V N											
AH 6-1289	V P T D H G A H L S L R G L P V C A F S S A G P C A L R F T S A R R M E T T V N .A.											
AH 6-2394	V P T D H G A H L S L R G L P V C A F S S A G P C A L R F T S A R R M E T T V N .A. .G.											
AH 6-2616	V P T D H G A H L S L R G L P V C A F S S A G P C A L R F T S A R R M E T T V N											
AH 7-1010	V P T D H G A H L S L R G L P V C A F S S A G P C A L R F T S A R R M E T T V N											
AH 7-1440	V P T D H G A H L S L R G L P V C A F S S A G P C A L R F T S A R R M E T T V N											
AH 7-1275	V P T D H G A L L S L R G L P V C A F S S A G P C A L R F T S A R R M E T T V N .T.											
AH 7-1442	V P T D H G A H L S L R G L P V C A F S S A G P C A L R F T S A R R M E T T V N											
AH 9-296	V P T D H G A H L S L R G L P V C A F S S A G P C A L R F T S A R R M E T T V N											
AH 9-2772	V P T D H G A H L S L R G L P V C A F S S A G P C A L R F T S A R R M E T T V N											
AH 10-2085	V P T D H G A H L S L R G L P V C A F S S A G P C A L R F T S A R R M E T T V N											
AH 11-2759	V P T D H G A H L S L R G L P V C A F S S A G P C A L R F T S A R R M E T T V N											
AH 11-3705	V S T D H G A H L S L R G L P V C A F S S A G P C A L R F T S A R R M E T T V N .T.											
AH 12-5311	V S T D H G A H L S L R G L P V C A F S S A G P C A L R F T S A R R M E T T V N .T.											
AH 12-8090	V S T D H G A H L S L R G L P V C A F S S A G P C A L R F T S A R R M E T T V N .T.											
CH 1-1546	V P T D H G A H L S L R G L P V C A F S S A G P C A L R F T S A R R M E T T V N											
CH 7-1117A	V P T D H G A H L S L R G L P V C A F S S A G P C A L R F T S A R R M E T T V N .C.											
CH 12-70A	V P T D H G A H L S L R G L P V C A F S S A G P C A L R F T S A R R M E T T V N											
CH 12-1133	V P T D H G A H L S L R G L P V C A F S S A G P C A L R F T S A R R M E T T V N											
COMPL 4-1725	V S T D H G A H L S L R G L P V C A F S S A G P C A L R F T S A R R M E T T V N .T.C. .T.											
COMPL 7-2083	V P T D H G A H L S L R G L P V C A F S S A G P C A L R F T S A R R M E T T V N .A.											
COMPL 11-3179	V P T D H G A H L S L R G L P V C A F S S A G P C A L R F T S A R R M E T T V N											
X75657	GTTCCAGCCGACCACGGGGCGCACCTCTCTTTACGCGGTCTCCCGTCTGTGCCTTCTCATCTGCCGGTCCGGTGTGCACCTTCGCTTCACCTCTGCACGTTGCATGGAGACCACCGTGAAC											
AH 2-10101	V P A D H G A H L S L R G L P V C S F S S A G P C A L R F T S A R R M E T T V N .T.GT. .A. .G.											
	V S S D H G A H L S L R G L P V C A F S S A G P C A L R F T S A R R M E T T V N											

	1630	1640	1650	1660	1670	1680	1690	1700	1710	1720	1730	1740
X02763	GCCCATCAGATCCTGCCCAAGGCTTACATAAGAGGACTCTGGACTCCAGCAATGTCAACGACCGACCTTGAGGCCTACTTCAAAGACTGTGTGTTAAAGACTGGGAGGACTGGGG											
AH 4-2015	A H Q I L P K V L H K R T L G L P A M S T T D L E A Y F K D C V F K D W E E L GT.....											
AH 7-640	A H Q I L P K V L H K R T L G L P A M S T T D L E A Y F K D C V F K D W E E L GA.....											
AH 8-1993	A H Q I L P K V L H K R T L G L P A M S T T D L E A Y F K D C V F K D W E E L G											
CH 5-2628	A H Q I L P K V L H K R T L G L P A M S T T D L E A Y F K D C V F K D W E E L G											
CH 12-2805T.....											
COMPL 4-2971	A H Q I L P K V L H K R T L G L P A M S T T D L E A Y F K D C V F K D W E E L G											
V01460REV	GCCCAAAATATGCCCAAGGCTTACATAAGAGGACTCTGGACTCTCAGCAATGTCAACGACCGACCTTGAGGCATACTCAAAGACTGTTGTTAAAGACTGGGAGGACTGGGG											
X02496	A H Q I L P K V L H K R T L G L S A M S T T D L E A Y F K D C L F K D W E E L GA.C.T.C.....T.T.....											
AH 2-2658	A O P F L P K V L H K R T L G L S V M S T T D L E A Y F K D C L F K D W E E L GT.C.....T.T.....											
AH 3-7755	A H Q F L P K V L H K R T L G L S V M S T T D L E A Y F K D C L F K D W E E L GT.C.....T.T.....											
AH 3-7991	A H Q F L P K V L H K R T L G L S V M S T T D L E A Y F K D C L F K D W E E L GC.....T.....G.A.....											
AH 3-987	A H Q I L P K V L H K R T L G L S A M S T T D L E A Y F K D C L F K D W E E L GT.C.....T.T.....											
AH 5-556	A H Q F L P K V L H K R T L G L S V M S T T D L E A Y F K D C L F K D W E E L GT.C.....T.T.....											
AH 5-920	A H Q F L P K V L H K R T L G L S V M S T T D L E A Y F K D C L F K D W E E L GT.C.....T.....											
AH 6-1289	A H Q F L P K V L H K R T L G L S A M S T T D L E A Y F K D C L F K D W E E L GT.C.....T.....											
AH 6-2394	A H Q F L P K V L H K R T L G L S A M S T T D L E A Y F K D C L F K D W E E L GT.C.....T.....											
AH 6-2616	A H Q F L P K V L H K R T L G L S A M S T T D L E A Y F K D C L F K D W E E L GT.C.....T.....											
AH 7-1010	A H Q F L P K V L H K R T L G L S A M S T T D L E A Y F K D C L F K D W E E L GT.C.....T.T.....											
AH 7-1440	A H Q F L P K V L H K R T L G L S V M S T T D L E A Y F K D C L F K D W E E L GT.C.....T.T.....											
AH 7-1275	A H Q L L P K V L Y K R T L G L S A M S T T D L E A Y F K D C L F K D W E E L GC.....T.....T.....											
AH 7-1442	A H Q F L P K V L H K R T L G L S V M S T T D L E A Y F K D C L F K D W E E L GT.C.....T.T.....											
AH 9-296	A H Q F L P K V L H K R T L G L S V M S T T D L E A Y F K D C L F K D W E E L GT.C.....T.T.....											
AH 9-2772	A H Q L L P K V L Y K R T L G L S A M S T T D L E A Y F K D C L F K D W E E L GC.....T.....T.....											
AH 10-2085	A H Q F L P K V L H K R T L G L S V M S T T D L E A Y F K D C L F K D W E E L GT.C.....T.T.....											
AH 11-2759	A H Q F L P K V L H K R T L G L S V M S T T D L E A Y F K D C L F K D W E E L GT.C.....T.T.....											
AH 11-3705	A H Q F L P K V L H K R T L G L S V M S T T D L E A Y F K D C L F K D W E E L GT.C.....T.T.....GT.G.....											
AH 12-5311	A H Q F L P K V L H K R T L G L S V M S T T D L E A Y F K D C L F K D W E E L GT.C.....T.T.....											
AH 12-8090	A H Q F L P K V L H K R T L G L S V M S T T D L E A Y F K D C L F K D W E E L GT.C.....T.T.....											
CH 1-1546	A H Q I L P K V L H K R T L G L S A M S T T D L E A Y F K D C L F K D W E E L GC.....T.....											
CH 7-1117A	A H Q I L P K V L H K R T L G L S A M S T T D L E A Y F K D C L F K D W E E L GC.....T.....G.....											
CH 12-70A	A H H F L P K V L H K R T L G L S A M S T T D L E A Y F K D C L F K D W E E L GCT.C.....T.TG.....											
CH 12-1133	A H Q F L P K V L H K R T L G L S V M S T T D L E A Y F K D C L F K D W E E L GT.C.....T.T.....C.....											
COMPL 4-1725	A H Q F L P K V L H K R T L G L S V M S T T D L E A Y F K D C L F K D W E E L GT.C.....C.....T.TG.....											
COMPL 7-2083	A H Q F L P K V L H K R T L G L S A M S T T D L E A Y F K D C L F K D W E E L GT.C.....T.....											
COMPL 11-3179	A H Q F L P K V L H K R T L G L S V M S T T D L E A Y F K D C L F K D W E E L GT.C.....T.T.....											

	1750	1760	1770	1780	1790	1800	1810	1820
X02763	GAGGAGATTAGGTTAAAGGCTCTTTGTATTAGGAGGCTGTAGGCACAAATTGGTCTGCGCACCCAGCACCATGCAACTTTTT							
AH 4-2015	E E I R L K V F V L G G C R H K L V C A P A P C N F	.T.A.						
AH 7-640	E E I R L M I F V L G G C R H K L V C A P A P C N F	.T.						
AH 8-1993	E E I R L K V F V L G G C R H K L V C A P A P C N F	.T.						
CH 5-2628	E E I R L K V F V L G G C R H K L V C A P A P C N F	.T.						
CH 12-2805	E E I R L K V F V L G G C R H K L V C A P A P C N F	.T.						
COMPL 4-2971	E E I R L K V F V L G G C R H K L V C A P A P C N F	.T.						
V01460REV	GAGGAGATTAGGTTAAAGGCTCTTTGTACTAGGAGGCTGTAGGCATAAATTGGTCTGCGCACCCAGCACCATGCAACTTTTT							
X02496	E E I R L K V F V L G G C R H K L V C A P A P C N F	.A.						
AH 2-2658	E E T R L K V F V L G G C R H K L V C A P A P C N F	.C.						
AH 3-7755	E E S R L M I F V L G G C R H K L V C A P A P C N F	.G..A..T.A.T.						
AH 3-7991	E E T R L M I F V L G G C R H K L V C A P A P C N F	.C..A..T.A.						
AH 3-987	E E I R L K V F V L G G C R H K L V C A P A P C N F	.A.						
AH 5-556	E E I R L K V F V L G G C R H K L V C A P A P C N F	.C.						
AH 5-920	E E T R L K V F V L G G C R H K L V C A P A P C N F	.C..A.						
AH 6-1289	E E L R L K V F V L G G C R H K L V C A P A P C N F	.C..A.						
AH 6-2394	E E L R L K V F V L G G C R H K L V C A P A P C N F	.C..A.						
AH 6-2616	E E I R L K V F V L G G C R H K L V C A P A P C N F	.A.						
AH 7-1010	E E I R L K V F V L G G C R H K L V C A P A P C N F	.CG.						
AH 7-1440	E E T R L K V F V L G G C R H K L V C A P A P C N F	.A..G.						
AH 7-1275	E E T R L R I F V L G G C R H K L V C A P A P C N F	.C..A..G.A..T.						
AH 7-1442	E E L R L K V F V L G G C R H K L V C A P A P C N F	.C..A.						
AH 9-296	E E T R L K I F V L G G C R H K L V C A P A P C N F	.C..A..A..T.						
AH 9-2772	E E T R L M I F V L G G C R H K L V C A P A P C N F	.C..A..T.A.						
AH 10-2085	E E I R L K V F V L G G C R H K L V C A P A P C N F	.A..T.						
AH 11-2759	E E I R L K V F V L G G C R H K L V C A P A P C N F	.A.						
AH 11-3705	E E I R L K V F V L G G C R H K L V C A P A P C N F	.C..A.						
AH 12-5311	E E L R L K V F V L G G C R H K L V C A P A P C N F	.C..A.						
AH 12-8090	E E I R L K V F V L G G C R H K L V C A P A P C N F	.A.						
CH 1-1546	E E T R L M I F V L G G C R H K L V C A P A P C N F	.C..A..T.A.						
CH 7-1117A	E E L R L K V F V L G G C R H K L V C A P A P C N F	.C..A.						
CH 12-70A	E E I R L K V F V L G G C R H K L V C A P A P C N F	.A.						
CH 12-1133	E E I R L K L F V L G G C R H K L V C A P T P C N F	.A..T.G..A.						
COMPL 4-1725	E E I R L K V F V L G G C R H K L V C A P A P C N F	.A.						
COMPL 7-2083	E E L R L K V F V L G G C R H K L V C A P A P C N F	.C..A.						
COMPL 11-3179	E E I R L K V F V L G G C R H K L V C A P A P C N F	.A.						

REFERENCES

[1] Arauz-Ruiz P, Norder H, Robertson BH, Magnus LO. Genotype H: a new Amerindian genotype of hepatitis B virus revealed in Central America. *J Gen Virol* 2002; 83(Pt 8): 2059-73.

[2] Norder H, Courouce AM, Magnus LO. Complete genomes, phylogenetic relatedness, and structural proteins of six strains of the hepatitis B virus, four of which represent two new genotypes. *Virology* 1994; 198(2): 489-503.

[3] Stuyver L, De GS, Van GC, *et al.* A new genotype of hepatitis B virus: complete genome and phylogenetic relatedness. *J Gen Virol* 2000; 81(Pt 1): 67-74.

[4] Kramvis A, Weitzmann L, Owiredu WK, Kew MC. Analysis of the complete genome of subgroup A' hepatitis B virus isolates from South Africa. *J Gen Virol* 2002; 83(Pt 4): 835-9.

[5] Kurbanov F, Tanaka Y, Fujiwara K, *et al.* A new subtype (subgenotype) Ac (A3) of hepatitis B virus and recombination between genotypes A and E in Cameroon. *J Gen Virol* 2005; 86(Pt 7): 2047-56.

[6] Norder H, Courouce AM, Coursaget P, *et al.* Genetic diversity of hepatitis B virus strains derived worldwide: genotypes, subgenotypes, and HBsAg subtypes. *Intervirology* 2004; 47(6): 289-309.

[7] Magnus LO, Norder H. Subtypes, genotypes and molecular epidemiology of the hepatitis B virus as reflected by sequence variability of the S-gene. *Intervirology* 1995; 38(1-2): 24-34.

[8] Chu CJ, Lok AS. Clinical significance of hepatitis B virus genotypes. *Hepatology* 2002; 35(5): 1274-6.

[9] Courouce-Pauty AM, Plancon A, Soulier JP. Distribution of HBsAg subtypes in the world. *Vox Sang* 1983; 44(4): 197-211.

[10] Pumpen PP, Dishler AV, Kozlovskaja TM, Bychko VV, Gren EI. Cloning of hepatitis B virus DNA in *Escherichia coli*. *Dokl Akad Nauk SSSR* 1981; 260(4): 1022-4.

- [11] Pumpen PP, Kozlovskaja TM, Dishler AV, Bychko VV, Kalis I. The comparative mapping of virion and cloned DNA of the hepatitis B virus. *Mol Biol (Mosk)* 1982; 16(6): 1314-21.
- [12] Dumpis U, Kovalova Z, Jansons J, *et al.* An outbreak of HBV and HCV infection in a paediatric oncology ward: epidemiological investigations and prevention of further spread. *J Med Virol* 2003; 69(3): 331-8.
- [13] Sominskaya I, Mihailova M, Jansons J, *et al.* Hepatitis B and C virus variants in long-term immunosuppressed renal transplant patients in Latvia. *Intervirology* 2005; 48(2-3): 192-200.
- [14] Preikschat P, Meisel H, Will H, Gunther S. Hepatitis B virus genomes from long-term immunosuppressed virus carriers are modified by specific mutations in several regions. *J Gen Virol* 1999; 80 (Pt 10): 2685-91.
- [15] Gunther S. Genetic variation in HBV infection: genotypes and mutants. *J Clin Virol* 2006; 36 (Suppl 1): S3-S11.
- [16] Yotsuyanagi H, Hino K, Tomita E, Toyoda J, Yasuda K, Iino S. Precore and core promoter mutations, hepatitis B virus DNA levels and progressive liver injury in chronic hepatitis B. *J Hepatol* 2002; 37(3): 355-63.
- [17] Takahashi K, Ohta Y, Kanai K, *et al.* Clinical implications of mutations C-to-T1653 and T-to-C/A/G1753 of hepatitis B virus genotype C genome in chronic liver disease. *Arch Virol* 1999; 144(7): 1299-308.
- [18] Lindh M, Hannoun C, Dhillon AP, Norkrans G, Horal P. Core promoter mutations and genotypes in relation to viral replication and liver damage in East Asian hepatitis B virus carriers. *J Infect Dis* 1999; 179(4): 775-82.
- [19] Brunetto MR, Giarin MM, Oliveri F, *et al.* Wild-type and e antigen-minus hepatitis B viruses and course of chronic hepatitis. *Proc Natl Acad Sci USA* 1991; 88(10): 4186-90.
- [20] Parvez MK, Thakur V, Kazim SN, Guptan RC, Hasnain SE, Sarin SK. Base-pair alterations in the epsilon-lower stem due to a novel double substitution in the precore gene of HBV-e negative variant were recovered by secondary mutations. *Virus Genes* 2001; 23(3): 315-20.
- [21] Chen J, Zhao WD, Wang M, Liu FG, Li JM. Effects of ethanol on HBV replication and gene expression in HBV transgenic mice. *Zhonghua Gan Zang Bing Za Zhi* 2005; 13(6): 471-2.
- [22] Ehata T, Omata M, Yokosuka O, Hosoda K, Ohto M. Variations in codons 84-101 in the core nucleotide sequence correlate with hepatocellular injury in chronic hepatitis B virus infection. *J Clin Invest* 1992; 89(1): 332-8.
- [23] Asahina Y, Enomoto N, Ogura Y, *et al.* Complete nucleotide sequences of hepatitis B virus genomes associated with epidemic fulminant hepatitis. *J Med Virol* 1996; 48(2): 171-8.
- [24] Aye TT, Uchida T, Becker SO, Shikata T, Mima S. Completely or nearly identical hepatitis B virus strains replicate between patients with acute or fulminant hepatitis B and their respective infectious sources. *J Med Virol* 1994; 42(1): 60-5.
- [25] Aye TT, Uchida T, Becker SO, *et al.* Variations of hepatitis B virus precore/core gene sequence in acute and fulminant hepatitis B. *Dig Dis Sci* 1994; 39(6): 1281-7.
- [26] Asahina Y, Enomoto N, Ogura Y, *et al.* Sequential changes in full-length genomes of hepatitis B virus accompanying acute exacerbation of chronic hepatitis B. *J Hepatol* 1996; 25(6): 787-94.
- [27] Hur GM, Lee YI, Suh DJ, Lee JH, Lee YI. Gradual accumulation of mutations in precore core region of HBV in patients with chronic active hepatitis: implications of clustering changes in a small region of the HBV core region. *J Med Virol* 1996; 48(1): 38-46.
- [28] Hong P, Ng LF, Ren EC, Chen WN. A cell-based system for hepatitis B virus replication: significance of clinically enhanced viral replication in relation to deletions in viral core promoter. *Front Biosci* 2005; 10: 2001-4.
- [29] Yamashita K, Kagawa S, Matsuoka A, Takarada Y. Detection of hepatitis B virus by using polymerase chain reaction and nonradioactive DNA probes. II. Identification of mutations in the core gene by PCR-direct sequencing and ASO probe method. *Kansenshogaku Zasshi* 1996; 70(10): 1072-8.
- [30] Bertoletti A, Costanzo A, Chisari FV, *et al.* Cytotoxic T lymphocyte response to a wild type hepatitis B virus epitope in patients chronically infected by variant viruses carrying substitutions within the epitope. *J Exp Med* 1994; 180(3): 933-43.
- [31] Pumpens P, Grens E, Nassal M. Molecular epidemiology and immunology of hepatitis B virus infection - an update. *Intervirology* 2002; 45(4-6): 218-32.
- [32] Okumura A, Takayanagi M, Aiyama T, *et al.* Serial analysis of hepatitis B virus core nucleotide sequence of patients with acute exacerbation during chronic infection. *J Med Virol* 1996; 49(2): 103-9.
- [33] Okumura A, Ishikawa T, Yoshioka K, Yuasa R, Fukuzawa Y, Kakumu S. Mutation at codon 130 in hepatitis B virus (HBV) core region increases markedly during acute exacerbation of hepatitis in chronic HBV carriers. *J Gastroenterol* 2001; 36(2): 103-10.
- [34] Takahashi K, Akahane Y, Hino K, Ohta Y, Mishihiro S. Hepatitis B virus genomic sequence in the circulation of hepatocellular carcinoma patients: comparative analysis of 40 full-length isolates. *Arch Virol* 1998; 143(12): 2313-26.
- [35] Webster GJ, Reignat S, Maini MK, *et al.* Incubation phase of acute hepatitis B in man: dynamic of cellular immune mechanisms. *Hepatology* 2000; 32(5): 1117-24.
- [36] Carman WF, Boner W, Fattovich G, *et al.* Hepatitis B virus core protein mutations are concentrated in B cell epitopes in progressive disease and in T helper cell epitopes during clinical remission. *J Infect Dis* 1997; 175(5): 1093-100.
- [37] Preikschat P, Gunther S, Reinhold S, *et al.* Complex HBV populations with mutations in core promoter, C gene, and pre-S region are associated with development of cirrhosis in long-term renal transplant recipients. *Hepatology* 2002; 35(2): 466-77.
- [38] Tai PC, Banik D, Lin GI, *et al.* Novel and frequent mutations of hepatitis B virus coincide with a major histocompatibility complex class I-restricted T-cell epitope of the surface antigen. *J Virol* 1997; 71(6): 4852-6.
- [39] Weinberger KM, Bauer T, Bohm S, Jilg W. High genetic variability of the group-specific a-determinant of hepatitis B virus surface antigen (HBsAg) and the corresponding fragment of the viral polymerase in chronic virus carriers lacking detectable HBsAg in serum. *J Gen Virol* 2000; 81(Pt 5): 1165-74.
- [40] Alexopoulou A, Baltayiannis G, Jammeh S, Waters J, Dourakis SP, Karayiannis P. Hepatitis B surface antigen variant with multiple mutations in the a determinant in an agammaglobulinemic patient. *J Clin Microbiol* 2004; 42(6): 2861-5.
- [41] Chen WN, Oon CJ, Lim GK. Hepatitis B virus surface antigen mutants persist in chronic carriers receiving lamivudine therapy in Singapore. *J Antimicrob Chemother* 2002; 49(6): 1044-6.
- [42] Bielawski KP, Dybikowska A, Lisowska-Charmuszko U, *et al.* Distribution of HBV genotypes and mutants among hepatitis B infected patients from northern Poland. *Int J Mol Med* 2004; 14(2): 301-4.
- [43] Tallo T, Norder H, Tefanova V, *et al.* Hepatitis B virus genotype D strains from Estonia share sequence similarity with strains from Siberia and may specify ayw4. *J Med Virol* 2004; 74(2): 221-7.
- [44] Olingier CM, Lazouskaya N, Erimin V, Muller CP. Molecular phylogenic of hepatitis B and C in Belarus. The International Meeting on the Molecular Biology of Hepatitis B and C Viruses 2005; pp. 157.
- [45] Grannikova SA, Rusakova EV, Budnitskaia PZ, Koroleva VM, Lorinskaia TM. Subtype characteristics of the surface antigen of hepatitis B in different regions of the USSR. *Zh Mikrobiol Epidemiol Immunobiol* 1977; (10): 97-101.
- [46] Abe K, Hayakawa E, Sminov AV, *et al.* Molecular epidemiology of hepatitis B, C, D and E viruses among children in Moscow, Russia. *J Clin Virol* 2004; 30(1): 57-61.
- [47] Netesova IG, Swenson PD, Osipova LP, Gubina MA, Posukh OL, Netesov SV. Determination of HBsAg subtypes in Western Siberian part of Russia. *J Med Virol* 2003; 71(2): 183-7.
- [48] Szomor KN, Dencs A, Toth G, *et al.* Variability of the PreS1/PreS2/S regions of hepatitis B virus in Hungary. *Arch Virol* 2007; 152(4): 697-704.
- [49] Dal MG, Poli A, Croce LS, *et al.* Hepatitis B virus genotypes, core promoter variants, and precore stop codon variants in patients infected chronically in North-Eastern Italy. *J Med Virol* 2006; 78(6): 734-40.
- [50] De MC, Giambelli C, Tanzi E, *et al.* High level of genetic heterogeneity in S and P genes of genotype D hepatitis B virus. *Virology* 2007; 365(1): 113-24.
- [51] Echevarria JM, Avellon A, Magnius LO. Molecular epidemiology of hepatitis B virus in Spain: identification of viral genotypes and

- prediction of antigenic subtypes by limited sequencing. J Med Virol 2005; 76(2): 176-84.
- [52] Jardi R, Rodriguez F, Buti M, *et al.* Mutations in the basic core promoter region of hepatitis B virus. Relationship with precore variants and HBV genotypes in a Spanish population of HBV carriers. J Hepatol 2004; 40(3): 507-14.
- [53] Micallessi MI, De CL, Vranckx R. Hepatitis B virus (HBV) genotyping in Belgian patients with chronic HBV infection. Clin Microbiol Infect 2005; 11(6): 499-501.
- [54] van Houdt R, Bruisten SM, Koedijk FD, *et al.* Molecular epidemiology of acute hepatitis B in the Netherlands in 2004: nationwide survey. J Med Virol 2007; 79(7): 895-901.

Received: June 24, 2010

Revised: August 14, 2010

Accepted: December 02, 2010

© Sominskaya *et al.*; Licensee Bentham Open.

This is an open access article licensed under the terms of the Creative Commons Attribution Non-Commercial License (<http://creativecommons.org/licenses/by-nc/3.0/>), which permits unrestricted, non-commercial use, distribution and reproduction in any medium, provided the work is properly cited.