

(19)日本国特許庁 (J P)

(12) 公表特許公報 (A)

(11)特許出願公表番号

特表2003 - 525428

(P2003 - 525428A)

(43)公表日 平成15年8月26日(2003.8.26)

(51)Int.Cl ⁷	識別記号	F I	テ-マコード* (参考)
G 0 1 N 33/569		G 0 1 N 33/569	L 4 B 0 2 4
A 6 1 K 35/14		A 6 1 K 35/14	Z 4 B 0 6 3
39/00		39/00	H 4 B 0 6 5
39/395		39/395	D 4 C 0 8 5
			S 4 C 0 8 7

審査請求 未請求 予備審査請求 (全279数) 最終頁に続く

(21)出願番号 特願2001 - 547552(P2001 - 547552)

(86)(22)出願日 平成12年12月19日(2000.12.19)

(85)翻訳文提出日 平成14年6月20日(2002.6.20)

(86)国際出願番号 PCT/US00/34420

(87)国際公開番号 W001/046696

(87)国際公開日 平成13年6月28日(2001.6.28)

(31)優先権主張番号 09/468,147

(32)優先日 平成11年12月21日(1999.12.21)

(33)優先権主張国 米国(US)

(71)出願人 アボット・ラボラトリーズ
ABBOTT LABORATORIE
S

アメリカ合衆国、イリノイ・60064 - 6050、
アボット・パーク、アボット・パーク・ロ
ード・100、チャド・0377/エイ・ピー・6・
デー2

(72)発明者 シュロイダー , ジョージ・ジー
アメリカ合衆国、イリノイ・60076、スコ
ーキー、カーロブ・7640

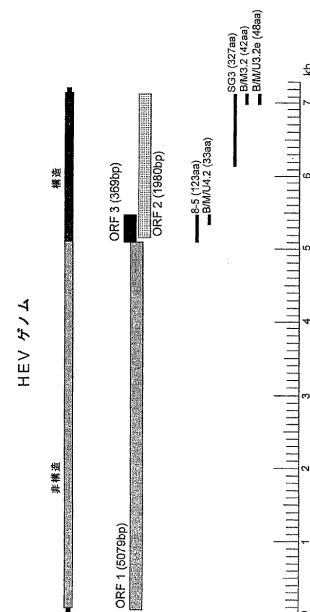
(74)代理人 弁理士 川口 義雄 (外 4 名)

最終頁に続く

(54)【発明の名称】 E 型肝炎ウイルスを検出するための方法及び組成物

(57)【要約】

サンプル中のU S タイプもしくはU S サブタイプE 型肝炎ウイルス (その天然変異体をも含む) の存在を検出するための方法及び組成物を開示する。特に、本発明はU S タイプもしくはU S サブタイプE 型肝炎ウイルスのゲノムに相当する核酸配列、前記ウイルスのゲノムによりコードされるアミノ酸配列 (エピトープ配列も含む) 、及び前記アミノ酸配列に特異的に結合する抗体を提供する。本発明は更に、前記ウイルスによる感染に対して患者を免疫化するためまたは前記ウイルスに既に感染した患者を治療するための方法及び組成物を提供する。



【特許請求の範囲】

【請求項1】 試験サンプル中のUSタイプもしくはUSサブタイプE型肝炎ウイルス（HEV）またはその天然変異体の存在を検出する方法であって、

（a）前記サンプル中に存在すると結合パートナーに結合してマーカー-結合パートナー複合体を生成する前記ウイルスに対するマーカーに特異的に結合する結合パートナーと前記サンプルを接触させるステップ、及び

（b）前記サンプル中の前記ウイルスの存在の指標である前記複合体の存在を検出するステップ

を含むことを特徴とする前記方法。

【請求項2】 マーカーがウイルスに結合し得る抗体であることを特徴とする請求の範囲第1項に記載の方法。

【請求項3】 抗体が、免疫グロブリンGまたは免疫グロブリンMであることを特徴とする、請求の範囲第2項に記載の方法。

【請求項4】 結合パートナーが単離ポリペプチド鎖であることを特徴とする、請求の範囲第2項に記載の方法。

【請求項5】 ポリペプチド鎖が固体支持体上に固定化されていることを特徴とする、請求の範囲第4項に記載の方法。

【請求項6】 結合パートナーが、天然変異体を含めた配列番号91、配列番号92及び配列番号93からなる群から選択されるポリペプチド鎖であることを特徴とする、請求の範囲第4項に記載の方法。

【請求項7】 結合パートナーが、配列番号173または配列番号175に示すアミノ酸配列を含むポリペプチド鎖であることを特徴とする、請求の範囲第4項に記載の方法。

【請求項8】 結合パートナーが、配列番号174または配列番号176に示すアミノ酸配列を含むポリペプチド鎖であることを特徴とする、請求の範囲第4項に記載の方法。

【請求項9】 結合パートナーが、それらの天然変異体を含めた配列番号166、配列番号167及び配列番号168からなる群から選択されるポリペプチド鎖であることを特徴とする、請求の範囲第4項に記載の方法。

【請求項10】 結合パートナーが、配列番号223に示すアミノ酸配列を含むポリペプチド鎖であることを特徴とする、請求の範囲第4項に記載の方法。

【請求項11】 結合パートナーが、配列番号224に示すアミノ酸配列を含むポリペプチド鎖であることを特徴とする、請求の範囲第4項に記載の方法。

【請求項12】 結合パートナーが、天然変異体を含めた配列番号91、配列番号92、配列番号93、配列番号166、配列番号167及び配列番号168からなる群から選択されるポリペプチド鎖に特異的に結合し得る単離抗体であることを特徴とする、請求の範囲第1項に記載の方法。

【請求項13】 抗体がモノクローナル抗体であることを特徴とする、請求の範囲第12項に記載の方法。

【請求項14】 マーカーがポリペプチド鎖であることを特徴とする、請求の範囲第1項に記載の方法。

【請求項15】 ポリペプチド鎖が、天然変異体を含めた配列番号91、配列番号92及び配列番号93からなる群から選択されることを特徴とする、請求の範囲第14項に記載の方法。

【請求項16】 ポリペプチド鎖が、配列番号173または配列番号175に示すアミノ酸配列を含むことを特徴とする、請求の範囲第14項に記載の方法。

【請求項17】 ポリペプチド鎖が、配列番号174または配列番号176に示すアミノ酸配列を含むことを特徴とする、請求の範囲第14項に記載の方法。

【請求項18】 ポリペプチド鎖が、天然変異体を含めた配列番号166、配列番号167及び配列番号168からなる群から選択されることを特徴とする、請求の範囲第14項に記載の方法。

【請求項19】 ポリペプチド鎖が配列番号223に示すアミノ酸配列を含むことを特徴とする、請求の範囲第14項に記載の方法。

【請求項20】 ポリペプチド鎖が配列番号224に示すアミノ酸配列を含むことを特徴とする、請求の範囲第14項に記載の方法。

【請求項21】 マーカーが、ウイルスのゲノムの少なくとも一部を規定す

る核酸配列またはその相補鎖であることを特徴とする、請求の範囲第1項に記載の方法。

【請求項22】 結合パートナーが、特定のハイブリダイゼーション条件下で、配列番号89及び164に示す核酸配列にハイブリダイズし得る単離核酸配列であることを特徴とする、請求の範囲第1項に記載の方法。

【請求項23】 結合パートナーが、配列番号126、配列番号128、配列番号147、配列番号148、配列番号150、配列番号152、配列番号177、配列番号178、配列番号255、配列番号256、配列番号257及び配列番号258からなる群から選択されることを特徴とする、請求の範囲第1項に記載の方法。

【請求項24】 結合パートナーが単離ポリペプチド鎖であることを特徴とする、請求の範囲第1項に記載の方法。

【請求項25】 試験サンプルが、哺乳動物細胞系であることを特徴とする、請求の範囲第1項に記載の方法。

【請求項26】 哺乳動物細胞系が、ヒト胎児腎細胞系であることを特徴とする、請求の範囲第41項に記載の方法。

【請求項27】 試験サンプル中のE型肝炎ウイルス(HEV)の存在の検出方法であって、

(a) サンプル中に存在すると結合パートナーに結合してマーカ-結合パートナー複合体を生成する前記ウイルスに対するマーカ-に特異的に結合する配列番号126、配列番号128、配列番号147、配列番号148、配列番号150、配列番号152、配列番号177、配列番号178、配列番号255、配列番号256、配列番号257及び配列番号258からなる群から選択される結合パートナーと前記サンプルを接触させるステップ、及び

(b) 前記サンプル中の前記ウイルスの存在の指標である前記複合体の存在を検出するステップ

を含むことを特徴とする前記方法。

【請求項28】 配列番号173、配列番号174、配列番号175、配列番号176、配列番号223及び配列番号224に示すアミノ酸配列を含むこと

を特徴とする単離ポリペプチド鎖。

【請求項29】 USタイプもしくはUSサブタイプHEVのORF1配列によりコードされるポリペプチド、USタイプもしくはUSサブタイプHEVのORF2配列によりコードされるポリペプチド、及びUSタイプもしくはUSサブタイプHEVのORF3配列によりコードされるポリペプチドからなる群から選択されるポリペプチド鎖に特異的に結合し得ることを特徴とする単離抗体。

【請求項30】 配列番号173、配列番号175または配列番号224に示すアミノ酸配列を含むポリペプチド鎖に特異的に結合し得ることを特徴とする単離抗体。

【請求項31】 配列番号174、配列番号176または配列番号223に示すアミノ酸配列を含むポリペプチド鎖に特異的に結合し得ることを特徴とする単離抗体。

【請求項32】 同様の条件下で、配列番号169または配列番号171に示すアミノ酸配列を含むポリペプチド鎖に対して低アフィニティーを有することを特徴とする、請求の範囲第30項に記載の単離抗体。

【請求項33】 同様の条件下で、配列番号170または配列番号172に示すアミノ酸配列を含むポリペプチド鎖に対して低アフィニティーを有することを特徴とする、請求の範囲第31項に記載の単離抗体。

【請求項34】 更に、検出可能部分を含むことを特徴とする、請求の範囲第29項に記載の単離抗体。

【請求項35】 USタイプもしくはUSサブタイプE型肝炎ウイルスのORF1、ORF2またはORF3配列の少なくとも一部を規定することを特徴とする単離核酸配列またはその相補配列。

【請求項36】 特定ハイブリダイゼーション条件下で配列番号89及び配列番号164に示すヌクレオチド配列にハイブリダイズし得ることを特徴とする単離核酸配列。

【請求項37】 請求の範囲第35項に記載の単離核酸配列を含むことを特徴とするベクター。

【請求項38】 請求の範囲第37項に記載のベクターを含むことを特徴と

する宿主細胞。

【請求項39】 USタイプもしくはUSサブタイプHEVに対して哺乳動物を免疫化する方法であって、前記哺乳動物に対して請求の範囲第28項に記載のポリペプチドを前記USタイプもしくはUSサブタイプE型肝炎ウイルスに特異的に結合し得る抗体の産生を刺激するのに十分な量投与することを特徴とする前記方法。

【請求項40】 USタイプもしくはUSサブタイプHEV 1に対して哺乳動物を免疫化する方法であって、前記哺乳動物に対して請求の範囲第29項に記載の抗体を前記USタイプもしくはUSサブタイプE型肝炎ウイルスに対して該哺乳動物を免疫化するのに十分な量投与することを特徴とする前記方法。

【請求項41】 USタイプもしくはUSサブタイプHEV 1に対して哺乳動物を免疫化する方法であって、前記哺乳動物に対して請求の範囲第30項に記載の抗体を前記USタイプもしくはUSサブタイプE型肝炎ウイルスに対して該哺乳動物を免疫化するのに十分な量投与することを特徴とする前記方法。

【請求項42】 USタイプもしくはUSサブタイプHEV 1に対して哺乳動物を免疫化する方法であって、前記哺乳動物に対して請求の範囲第31項に記載の抗体を前記USタイプもしくはUSサブタイプE型肝炎ウイルスに対して該哺乳動物を免疫化するのに十分な量投与することを特徴とする前記方法。

【請求項43】 USタイプもしくはUSサブタイプHEVに対して哺乳動物を免疫化する方法であって、前記哺乳動物に対して請求の範囲第38項に記載の宿主細胞を前記USタイプもしくはUSサブタイプE型肝炎ウイルスに対して該哺乳動物を免疫化するのに十分な量投与することを特徴とする前記方法。

【発明の詳細な説明】

【0001】

(発明の分野)

本発明は、包括的にはE型肝炎ウイルスを検出するための方法及び組成物に関し、より具体的にはE型肝炎ウイルスのUSタイプ及びUSサブタイプ株に感染している患者を検出または治療するための方法及び組成物に関する。

【0002】

(発明の背景)

肝臓の炎症(肝炎)の原因である肝親和性ウイルスには少なくとも5つの主要クラスがある。これらのウイルスにはA型肝炎ウイルス(HAV)、B型肝炎ウイルス(HBV)、C型肝炎ウイルス(HCV)、D型肝炎ウイルス(HDV)及びE型肝炎ウイルス(HEV)が含まれる。HBV、HCV及びHDVのみが慢性肝炎を引き起こすが、5タイプすべてのウイルスが直接にまたは例えばHBVやHDVの重感染/共感染の結果として急性疾患を引き起こす。HEVは他のウイルスと同様の肝炎の症状、例えば腹痛、黄疸、倦怠感、食欲不振、濃尿、発熱、悪心及び嘔吐を発現する(例えば、Reyesら、「非A非B肝炎剤の分子生物学：C型及びE型肝炎ウイルス(Molecular biology of non-A, non-B hepatitis agents: hepatitis C and hepatitis E viruses)」, *Advances in Virus Research*, 40:57-102(1991)、Bradley, 「非A非B肝炎ウイルスはC型及びE型ウイルスとして同定されるようになる(Hepatitis non-A, non-B viruses become identified as hepatitis C and E viruses)」, *Progr. Med. Virol*, 37:101-135(1990)、Hollinger, 「非A非B肝炎ウイルス(Non-A, non-B hepatitis viruses)」, *Virology*, 第2版, ニューヨーク州に所在の Raven Press(1990年)発行, p.2239-2271、Gustら, 「研究会報告：水伝播性非A非B肝炎ウイルス(Report of a workshop: water borne non-A, non-B hepatitis)」, *J. Infect. Dis.*, 156:630-635(1987)及びKrawczynski, 「E型肝炎(Hepatitis E)」, *Hepatology*, 17:932-941(1993)参照)。し

かしながら、他の肝炎ウイルスとは異なり、H E Vは通常米国では肝炎の重大な原因であると考えられていない。

【0003】

H E Vが流行している地域は東アフリカ、北アフリカ、インド、パキスタン、ビルマ及び中国である（上掲のR e y s（1991））。H E V感染の死亡率はH E Vが流行している地域の全人口の約0.1～約1.0%であると推定され、発展途上国の妊娠女性では約20%くらい高いと推定されている。多くは劇症肝炎のために死亡している（上掲のR e y s（1991））。米国、西ヨーロッパ及び日本では、H E Vが流行している地域を訪問してから自宅に戻った旅行者のH E V感染例が時々報告されている。しかしながら、H E V感染が中央官庁に報告されていないので、米国ではH E V感染の罹患率及び／または死亡率に関する情報は殆どない。米国では、H E Vの重大さを調べる系統的研究は今まで実施されていない。更に、前記研究が実施されていたとしても、前記研究を実施するための適切な試薬が開発されない限り米国（及び多分日本及び西ヨーロッパ）でのH E Vの相対的重大さは過小評価され続ける恐れがある。

【0004】

H E Vの基本的特徴は、H E Vが当業界においてオープンリーディングフレーム1（O R F 1）、オープンリーディングフレーム2（O R F 2）及びオープンリーディングフレーム3（O R F 3）と称される3つの非連続オープンリーディングフレーム（O R F）を含むポジティブセンスな一本鎖RNAゲノムを有する直径約27～30nmのエンベロープを持たないウイルスであることである。ウイルスの全体形態並びにゲノムの大きさ及び構成に基づいて、ウイルスは暫定的にカリシウイルス科の1つとして分類されている。最初に同定され、配列決定されたH E Vの2つの単離物はビルマ及びメキシコから入手した。両単離物のゲノムの範囲の全核酸同一性は76%である（R e y e sら, S c i e n c e , 247:1335-1339（1990）、T a mら, V i r o l o g y , 185:120-131（1991）、H u a n gら, V i r o l o g y , 191:550-558（1992））。ヌクレオチドの違いの多くは第3コドン位置で見られ、H E Vのビルマ株とメキシコ株間のアミノ酸配列の推定類似性はオープンリ

ーディングフレームORF 1、ORF 2及びORF 3のそれぞれに対して83%、93%及び87%であった。

【0005】

ビルマ株ではゲノムの5'末端に約27ヌクレオチドの短い非翻訳領域があるが、これは、メキシコ株では同定されていなかった。ORF 1は、推定メチルトランスフェラーゼドメイン、RNAヘリカーゼドメイン、推定RNA依存性RNAポリメラーゼ(RDRP)ドメイン及び推定パピン様プロテアーゼを含めた幾つかの保存モチーフをコードする約5,100ヌクレオチドを含む。すべてのポジティブセンスなRNA植物及び動物ウイルスに見られるGly - Asp - Asp (GDD)のトリペプチド配列はORF 1内にあり、通常RDRP機能を示す。通常細胞ヘリカーゼ及びウイルスヘリカーゼに関連するプリンNTPase活性を暗示する保存モチーフもORF 1配列中に存在する。ORF 1においてコードされる遺伝子産物に対する不変の免疫応答はない。

【0006】

第2のオープンリーディングフレーム(ORF 2)はウイルスゲノムの1/3のカルボキシルを占めている。ORF 2は、ORF 2のアミノ末端のコンセンサス信号ペプチド配列及び推定カプシドタンパク質をコードする約2,000ヌクレオチドを含み、該ヌクレオチドはORF 1に関連して+1リーディングフレームに翻訳される。HEV感染患者はORF 2から誘導されるペプチドまたは組み換えタンパク質と反応する抗体を産生していることが多い。

【0007】

第3のオープンリーディングフレーム(ORF 3)は、ORF 1及びORF 2と部分的に重複しており、ORF 1に関連して+2リーディングフレームに翻訳される369ヌクレオチドを含む。ORF 3によりコードされるタンパク質の機能は未知であるが、タンパク質は抗原性であり、多くのHEV感染患者がこのタンパク質に対する抗体を産生している。従って、ORF 2及びORF 3から誘導されるペプチドまたは組み換えタンパク質はHEVへの接触を診断するのに有用な血清学的マーカーとして使用し得る。

【0008】

最近、幾つかの別のH E V単離物が同定され、H E Vのビルマ株及びメキシコ株と比較された。最近の単離物の多くはH E Vのメキシコ株よりもビルマ株により密接に関連している。1986～1987年に短期間出現したものを除いて、H E Vのメキシコ株の単離物は更に存在しない(V e l a s q u e z ら, J A M A, 263:3281-3286(1992))。

【0009】

S A R - 55と称する1つの単離物が最近パキスタンのH E V感染患者から単離された。S A R - 55単離物はビルマ株と非常に関連しており、全ゲノムの範囲でのヌクレオチド及びアミノ酸同一性はそれぞれ94%及び99%である。他に幾つかの最近の単離物はパキスタンとの国境の中国のX u a r 省から単離された。これらの中国単離物は、ビルマ株(ヌクレオチド同一性約93%)よりもパキスタン株(ヌクレオチド同一性約98%)により密接に関連していた。

【0010】

ウイルスゲノムを配列決定したり、ウイルスコードされる組み換えタンパク質及び合成タンパク質を入手する前に、H E V感染を電子顕微鏡検査及び免疫蛍光法によりモニターした。H E Vゲノムが同定されたらすぐに、(i)特定のイムノアッセイ、例えば組み換えタンパク質及び/または合成ペプチドをベースとするウェスタンブロットアッセイ及びE L I S A、及び(ii)ポリメラーゼ連鎖反応(P C R)、例えば逆転写酵素P C R(R T - P C R)を含めたH E V感染を検出するための特定の検査方法が利用できるようになった。R T - P C Rを使用すると、急性肝炎感染の症例及びE T - N A N B Hが流行しているときに便または血清サンプル中のH E V R N Aがうまく検出された。更に、H E Vのメキシコ株及びビルマ株から誘導した組み換え抗原を使用することにより、ソマリア、ビルマ、ボルネオ、タシケント、ケニア、パキスタン及びメキシコのE T - N A N B H大発生から入手した試料においてH E Vに対する特異的I g G、I g M及び場合によりI g A抗体が検出された。エジプト、インド、タジキスタン及びウズベキスタンのような地域での急性散発性肝炎の症例及びH E Vが流行している地域に旅行した工業国(例えば、米国、英国、オランダ及び日本)の患者の中の急性肝炎の症例においてH E Vに対する特異的I g G及び場合によりI g M抗

体が検出された。

【0011】

今まで、H E Vのビルマ株及びメキシコ株単離物に対するP C R及びイムノアッセイに基づく試験で、“水伝播性肝炎”の各種症例がH E Vにより引き起こされることが確認された。抗体試験も発展途上国及びH E Vが流行している地域への旅行者の中での急性散発性肝炎の原因としてH E Vを確認するために重要であった。しかしながら、現在の試薬がH E Vのすべての株への接触を検出できないためにどれだけ多くの急性H E Vの症例が現在診断未確定になっているかは分からない。従って、H E Vの新しい単離物が同定されているので、現在市販されている試験キットでは今まで検出できなかった新しいH E V株により生ずる肝炎を検出及び／または治療するための新規な組成物及び方法の開発が望まれている。

【0012】

(発明の要旨)

本発明は、一部はヒトE型肝炎ウイルスの新規ファミリーの発見に基づく。新しく知見されたヒトE型肝炎ウイルスのファミリーは、以後U SタイプE型肝炎ウイルスと称するクラスに属する。更に、このファミリーの2つのメンバーが米国に居住している患者で発見され、ヌクレオチドレベル及びアミノ酸レベルで比較してかなりの類似性を示す。後者の2つのメンバーは共に、以後U SサブタイプE型肝炎ウイルスと称するU SタイプE型肝炎ウイルスのサブクラスに属する。

【0013】

従って、1つの態様で、本発明は当該試験サンプル中のU SタイプもしくはU SサブタイプE型肝炎ウイルスの存在を検出する方法を提供する。前記方法は、
(a) 前記サンプル中に存在すると結合パートナーに結合してマーカ－結合パートナー複合体を生成する前記ウイルスに対するマーカ－(または、標的)に特異的に結合する結合パートナーと前記サンプルを接触させるステップ、及び
(b) 前記複合体の存在または不在を検出するステップ
を含む。前記複合体の存在が試験サンプル中のウイルスの存在の指標である。

【0014】

1つの態様で、前記マーカーは当該サンプル中に存在するUSタイプもしくはUSサブタイプ抗体、例えば免疫グロブリンG (IgG) または免疫グロブリンM (IgM) であり、前記結合パートナーはマーカーに特異的に結合するエピトープを規定する単離ポリペプチド鎖である。この場合、試験サンプルは検査する患者から集めた体液サンプル、例えば血液、血清または血漿であると考えられる。好ましい実施態様では、USタイプもしくはUSサブタイプ特定エピトープを規定するポリペプチド鎖は固体支持体上に固定化されている。その後、固定化したポリペプチド鎖を、サンプル中に存在するマーカー抗体（例えば、抗-USタイプもしくは抗-USサブタイプE型肝炎ウイルス特異的抗体）が固定化ポリペプチドに結合し得る条件下でサンプルと混合する。その後、結合した抗体の存在または不在を、例えば検出可能部分で標識した第2抗体またはその抗原結合断片（例えば、抗-ヒト抗体またはその抗原結合断片）を用いて検出し得る。

【0015】

多種多様なUSタイプもしくはUSサブタイプ特異的ポリペプチドが本発明の上記実施態様を実施する際に結合パートナーとして有用であり得ると考えられる。例えば、本発明の1つの好ましい実施態様では、結合パートナーは天然変異体を含めた配列番号91、92及び93からなる群から選択され、ビルマファミリー及びメキシコファミリーのメンバーの対応アミノ酸配列と比較してユニークなアミノ酸配列を表すポリペプチド鎖の少なくとも一部、例えば少なくとも5個、好ましくは少なくとも8個、より好ましくは少なくとも15個、更に好ましくは少なくとも約25個のアミノ酸残基であり得ると考えられる。また、結合パートナーは配列番号173、174または175に示すアミノ酸配列を含むポリペプチド鎖であり得ると考えられる。本発明の別の好ましい実施態様では、結合パートナーは天然変異体を含めた配列番号166、167及び168からなる群から選択され、ビルマファミリー及びメキシコファミリーのメンバーの対応アミノ酸配列と比較してユニークなアミノ酸配列を表すポリペプチド鎖の少なくとも一部、例えば少なくとも5個、好ましくは少なくとも8個、より好ましくは少なくとも15個、更に好ましくは少なくとも約25個のアミノ酸残基であり得ると考えられる。また、結合パートナーは配列番号176、223または224に示すア

ミノ酸配列を含むポリペプチド鎖であり得ると考えられる。

【0016】

本発明の別の実施態様では、マーカーはHEVのUSタイプもしくはUSサブタイプファミリーのメンバーにとってユニークなポリペプチド鎖であり、結合パートナーはマーカーポリペプチド鎖上のエピトープに結合する単離抗体（例えば、ポリクローナルまたはモノクローナル抗体）であることが好ましい。結合パートナーは検出可能部分で標識されていても固体支持体上に固定化されていてもよい。例えば、固体支持体上に当該マーカーポリペプチド上の第1エピトープに結合する第1抗体を固定化することにより上記した本発明の実施態様は容易に実施され得ると考えられる。次いで、固定化抗体をマーカーポリペプチドに結合させ得る条件下で分析すべき試験サンプルを固体支持体と混合する。その後、結合したマーカーポリペプチド鎖の存在または不在を、例えばマーカーポリペプチド鎖上の第2の別のエピトープに結合する検出可能部分をコンジュゲートした第2抗体を用いて測定し得る。

【0017】

上記した本発明の実施態様を実施する際に有用な抗体は、好ましくは天然変異体を含めた配列番号91、92及び93からなる群から選択されるポリペプチド鎖に特異的に結合し得、ビルマファミリー及びメキシコファミリーのメンバーの対応配列と比較して前記ポリペプチド鎖に対してより高い結合アフィニティーを有することが好ましい。本発明を実施する際に有用な抗体は配列番号173または175に示すアミノ酸配列を含むポリペプチド鎖に特異的に結合し得ることが好ましいと考えられる。この抗体は更に、同様の条件下で、配列番号169または171に示すアミノ酸配列または配列番号175に相当するビルマ株及びメキシコ株中の領域に対して好ましくは低いアフィニティーを有するもの、最も好ましくは結合しないものとして特徴づけられる。また、本発明の実施の際に有用な抗体は好ましくは配列番号174または配列番号176に示すアミノ酸配列を含むポリペプチド鎖に特異的に結合し得ると考えられる。前記抗体は更に、同様の条件下で、配列番号170または172に示すアミノ酸配列または配列番号176に相当するビルマ株及びメキシコ株中の領域に対して好ましくは低いアフィニ

ティーを有するもの、最も好ましくは結合しないものとして特徴づけられる。

【0018】

また、上記した本発明の実施態様を実施する際に有用な抗体は天然変異体を含めた配列番号166、配列番号167及び配列番号168からなる群から選択されるポリペプチド鎖に特異的に結合し得、ビルマファミリー及びメキシコファミリーのメンバーの対応配列に比して前記ポリペプチド鎖に対して高い結合アフィニティーを有することが好ましいと考えられる。本発明の実施態様を実施する際に有用な抗体は配列番号223に示すアミノ酸配列を含むポリペプチド鎖に特異的に結合し得ることが好ましいと考えられる。この抗体は更に、同様の条件下で、配列番号170または172に示すアミノ酸配列に対して好ましくは低いアフィニティーを有するもの、最も好ましくは結合しないものとして特徴づけられる。また、本発明の実施態様を実施する際に有用な抗体は好ましくは配列番号224に示すアミノ酸配列を含むポリペプチド鎖に特異的に結合し得ると考えられる。この抗体は更に、同様の条件下で、配列番号169または171に示すアミノ酸配列に対して好ましくは低いアフィニティーを有するもの、最も好ましくは結合しないものとして特徴づけられる。

【0019】

本発明の別の実施態様では、マーカーはUSタイプもしくはUSサブタイプE型肝炎ウイルスのゲノムの少なくとも一部を規定する核酸配列またはその相補配列である。また、結合パートナーは例えば特定のハイブリダイゼーション条件または特定のPCRアニーリング条件下で配列番号89または164に示すヌクレオチド配列に特異的にハイブリダイズし得る、好ましくは8～100ヌクレオチド、より好ましくは10～75ヌクレオチド、最も好ましくは15～50ヌクレオチドを含む単離核酸配列（例えば、デオキシリボ核酸（DNA）、リボ核酸（RNA）またはペプチジル核酸（PNA）配列）であると考えられる。

【0020】

上記した本発明の実施態様は、例えば当該サンプルから核酸を単離することにより容易に実施され得る。その後、得られた核酸を例えばゲル電気泳動にかけ、固体支持体（例えば、ニトロセルロースまたはナイロン膜）に移し、固定化する

ことにより分画化しても、または慣用されているドットプロットまたはスロットプロット方法により固体支持体上に直接固定化してもよい。次いで、固定化した核酸を、マーカー配列に対して特異的にハイブリダイズし得る検出可能部分で標識した所定核酸配列を用いてプローブし得る。或いは、サンプル中のマーカー核酸の存在はポリメラーゼ連鎖反応（PCR）のような一般的な増幅方法により調べられ得、ここで特定の増幅産物の存在がサンプル中のマーカー核酸の存在の指標である。

【0021】

別の態様で、本発明は単離したUSタイプもしくはUSサブタイプ特異的ポリペプチド配列を提供する。これらのポリペプチドには、結合パートナーとして有用なUSタイプもしくはUSサブタイプE型肝炎特異的ポリペプチド鎖に関する欄で上記したものが含まれる。好ましい実施態様において、単離ポリペプチド鎖は配列番号93、配列番号168、配列番号173、配列番号174、配列番号175、配列番号176、配列番号223または配列番号224に示すアミノ酸配列を含む。上記した及び他のUSタイプもしくはUSサブタイプ特異的ポリペプチド鎖がサンプル中の抗-USタイプもしくは抗-USサブタイプE型肝炎特異的抗体の存在を検出するためのアッセイフォーマットにおいて使用され得ると考えられる。更に、前記ポリペプチドは実験動物において抗体を産生するために単独でまたはアジュバントと組合せて使用され得、または哺乳動物を予防または治療目的で免疫感作するためのワクチンとして医薬的に許容され得る担体と組合せて使用され得る。

【0022】

別の態様で、本発明は単離した抗-USタイプもしくは抗-USサブタイプE型肝炎特異的抗体を提供し、この中には結合パートナーとして有用な抗体に関する欄で上記したものが含まれる。好ましい実施態様において、単離抗体は、USタイプもしくはUSサブタイプE型肝炎ウイルスのORF1配列によりコードされるポリペプチド、USタイプもしくはUSサブタイプE型肝炎ウイルスのORF2配列によりコードされるポリペプチド、またはUSタイプもしくはUSサブタイプE型肝炎ウイルスのORF3配列によりコードされるポリペプチドからな

る群から選択されるポリペプチド鎖に特異的に結合し得る。特に、有用な抗体は配列番号93、配列番号168、配列番号173、配列番号174、配列番号175、配列番号176、配列番号223または配列番号224に示すアミノ酸配列を含むポリペプチド鎖に特異的に結合し得るという特徴を有すると考えられる。上記抗体及び他の抗体はサンプル中のUSタイプもしくはUSサブタイプE型肝炎ファミリーのメンバーの存在を検出するためのイムノアッセイにおいて有利に使用され得ると考えられる。抗体は、抗体自体が好ましくは検出可能部分で標識されてなる直接イムノアッセイにおいても、抗体自体が第2結合パートナーの標的、例えば検出可能部分で標識した第2抗体となる間接イムノアッセイにおいても使用され得る。更に、前記抗体は例えば哺乳動物を治療または予防目的で受動免疫するのに使用するために医薬的に許容され得る担体と組合せて使用され得ると考えられる。

【0023】

別の態様で、本発明は単離核酸配列、例えばサンプル中のUSタイプもしくはUSサブタイプE型肝炎ウイルスの存在を検出するためのマーカーまたは結合パートナーとしての核酸の使用に関する欄に記載したものを提供する。好ましい実施態様において、本発明はUSタイプもしくはUSサブタイプE型肝炎ウイルスのORF1、ORF2またはORF3配列の少なくとも一部を規定する単離核酸配列またはその相補配列を提供する。上記した及び他の核酸配列は、例えば当該サンプル中のUSタイプもしくはUSサブタイプE型肝炎ウイルスの存在を検出するためのヌクレオチドプローブ及び/または増幅プライマーとして使用され得ると考えられる。更に、核酸配列またはその相補配列はアンチセンス治療に使用するための医薬的に許容され得る担体と組合せられ得ると考えられる。更に、核酸配列はその後当該宿主細胞に形質転換またはトランスフェクトされるベクターに統合され得ると考えられる。その後、宿主細胞は医薬的に許容され得る担体と組合せて、哺乳動物を所定のUSタイプもしくはUSサブタイプE型肝炎に対して予防または治療目的で免疫化するためのワクチン（例えば、組み換えワクチン）として使用され得る。

【0024】

本発明の上記した及び他の目的、特徴及び作用効果は本発明の好ましい実施態様に関する以下の詳細説明から明らかとなるであろう。

【0025】

(図面の簡単な説明)

本発明の目的及び特徴は添付図面を参照することによりより深く理解され得る。

【0026】

図1は、ORF1、ORF2及びORF3領域の相対位置を示すHEVゲノムの概略図である。

【0027】

図2は、入院1日目から37日目までの患者USP-1中の血清アスパルテートアミノトランスフェラーゼ(四角)及び血清総ビリルビン(菱形)のレベルを示すグラフである。

【0028】

図3は、本研究中に単離したクローンの相対位置を示すHEV US-1ゲノムの概略図である。

【0029】

図4は、本研究中に単離したクローンの相対位置を示すHEV US-2ゲノムの概略図である。

【0030】

図5は、完全長HEV US-1、HEV US-2及び10個の他のHEV単離物由来のヌクレオチド配列の関係を示す根なし系統樹を示す。枝の長さは配列間の進化距離に比例する。位置あたりのヌクレオチド置換を示すスケールを示す。節間数(internal node numbers)は100反復から得た(すべての樹の%として表す)ブーツストラップ値を示す。示した単離物はビルマB1、B2；中国のC1、C2、C3、C4；パキスタンP1；インドI1、I2；メキシコM1；及び米国US-1、US-2である。

【0031】

図6は、ORF2/3領域由来のヌクレオチド配列(すなわち、配列番号89

のヌクレオチド残基5094～7114に相当する配列)の関係を示す根なし系統樹を示す。枝の長さは配列間の進化距離に比例する。位置あたりのヌクレオチド置換を示すスケールを示す。節間数は100反復から得た(すべての樹の%として表す)ブーツストラップ値を示す。示した単離物はビルマB1、B2;中国C1、C2、C3、C4;パキスタンP1;インドI1、I2;メキシコM1;ブタS1;及び米国US-1、US-2である。

【0032】

図7は、患者USP-2から採取した血清を接種する前またはその後のマカク中のアラニンアミノトランスフェラーゼ(四角)、血清アスパルテートトランスフェラーゼ(丸)及び γ -グルタミルトランスフェラーゼ(三角)のレベルを示すグラフである。HEV US-2 RNAが血清または糞サンプル中に存在するときの時間及び抗-HEV US-2 IgM及びIgGが検出可能な時間をも示す。

【0033】

図8は、本研究中に単離したクローンの相対位置を示すIt1ゲノム概略図である。

【0034】

図9は、ORF1、ORF2-3及びORF2に対するHEVコンセンサスプライマーの設計を示すビルマ(B1)、メキシコ(M1)、中国(C1)、パキスタン(P1)及びUS-1のアラインメントを示す。好ましいコンセンサスプライマーを強調して示す。

【0035】

図10は、371ヌクレオチド長で配列番号89の残基26～396に相当するORF1ヌクレオチド配列の関係を示す根なし系統樹を示す。位置あたりのヌクレオチド置換を示すスケールを示す。節間数は1000反復から得た(すべての樹の%として表す)ブーツストラップ値を示す。示した単離物はビルマB1、B2;中国C1、C2、C3、C4;パキスタンP1;インドI1、I2;メキシコM1;イタリアIt1;ギリシャG1、G2;及び米国US-1、US-2である。

【0036】

図11は、148ヌクレオチド長で配列番号89の残基6307～6454に相当するORF2ヌクレオチド配列の関係を示す根なし系統樹を示す。位置あたりのヌクレオチド置換を示すスケールを示す。節間数は1000反復から得た(すべての樹の%として表す)ブーツストラップ値を示す。示した単離物はビルマB1、B2；中国C1、C2、C3、C4；パキスタンP1；インドI1、I2；メキシコM1；イタリアIt1；ギリシャG1、G2；ブタS1；及び米国US-1、US-2である。

【0037】

図12は、好ましいHEV-US組み換えタンパク質構築物の概略図である。図12Aには、HEVのORF2及びORF3構造タンパク質を最初及び最後のアミノ酸位置と共に示す。免疫優性エピトープの存在をORFの範囲内のラインで示す。図12Bには、発現ベクターにクローン化したORF3領域を最初及び最後のアミノ酸位置と共に示す(配列番号203または204)。図12Cには、発現ベクターにクローン化したORF2領域を最初及び最後のアミノ酸位置と共に示す(配列番号199または200)。図12Dには、発現ベクターにクローン化したORF3/2キメラ構築物を該キメラ構築物の各成分の最初及び最後のアミノ酸位置と共に示す(配列番号206または207)。ORF3/2構築物から除去した配列を破線で示す。図12B～12Dでは、各構築物のカルボキシル末端のFLAG(登録商標)ペプチドの存在を黒四角で示す。

【0038】

図13は、患者USP-2から採取した血清を接種する前及びその後のマカク中のアラニンアミノトセンスフェラーゼ(四角)、IgG(円)及びIgM(星)のレベルを示すグラフである。

【0039】

図14は、371ヌクレオチド長で配列番号89の残基26～396に相当するORF1ヌクレオチド配列の関係を示す根なし系統樹を示す。位置あたりのヌクレオチド置換を示すスケールを示す。節間数は1000反復から得た(すべての樹の%として表す)ブーツストラップ値を示す。示した単離物はビルマB1、

B 2 ; 中国 C 1、C 2、C 3、C 4 ; パキスタン P 1 ; インド I 1、I 2 ; メキシコ M 1 ; イタリア I t 1 ; ギリシャ G 1、G 2 ; オーストリア A u 1 ; アルゼンチン A r 1、A r 2 ; 及び米国 U S - 1、U S - 2 である。

【0040】

図15は、148ヌクレオチド長で配列番号89の残基6307～6454に相当するORF2ヌクレオチド配列の関係を示す根なし系統樹を示す。位置あたりのヌクレオチド置換を示すスケールを示す。節間数は1000反復から得た(すべての樹の%として表す)ブーツストラップ値を示す。示した単離物はビルマ B 1、B 2 ; 中国 C 1、C 2、C 3、C 4 ; パキスタン P 1 ; インド I 1、I 2 ; メキシコ M 1 ; イタリア I t 1 ; ギリシャ G 1、G 2 ; オーストリア A u 1 ; アルゼンチン A r 1 ; ブタ S 1 ; 及び米国 U S - 1、U S - 2 である。

【0041】

図16は、98ヌクレオチド長で配列番号89の残基6354～6451に相当するORF2ヌクレオチド配列の関係を示す根なし系統樹を示す。位置あたりのヌクレオチド置換を示すスケールを示す。節間数は1000反復から得た(すべての樹の%として表す)ブーツストラップ値を示す。示した単離物はビルマ B 1、B 2 ; 中国 C 1、C 2、C 3、C 4 ; パキスタン P 1 ; インド I 1、I 2 ; メキシコ M 1 ; イタリア I t 1 ; ギリシャ G 1、G 2 ; オーストリア A u 1 ; アルゼンチン A r 1、A r 2 ; ブタ S 1 ; 及び米国 U S - 1、U S - 2 である。

【0042】

(発明の詳細説明)

上記したように、本発明の一部はヒトE型肝炎ウイルスの新しいファミリーの発見に基づく。E型肝炎ウイルスの新しく発見されたファミリーは、以後USタイプE型肝炎ウイルスと称するクラスに属する。更に、上記したように、USタイプのファミリーの2つのメンバーが米国に居住する二人の患者から得た血清において同定された。これら2つのメンバーは共に、以後USサブタイプE型肝炎ウイルスと称するUSタイプE型肝炎ウイルスのサブクラスに属する。USタイプ及びUSサブタイプE型肝炎ウイルスが発見されたことにより、従来市販されている肝炎検出キットを用いたときには肝炎を患っていると診断されなかった患

者においてUSタイプ及びUSサブタイプE型肝炎ウイルスの存在を検出するための方法及び組成物並びに前記ウイルスに対して患者を免疫化するための方法及び組成物の開発が可能となった。

【0043】

1つの態様で、本発明は試験サンプル中のUSタイプもしくはUSサブタイプE型肝炎ウイルスの存在の検出方法に関する。この方法は、(a)前記サンプル中に存在すると結合パートナーに結合してマーカー-結合パートナー複合体を生成する前記ウイルスに対す目マーカーに特異的に結合する結合パートナーと前記サンプルを接触させ、ステップ及び(b)前記複合体の存在を検出するステップを含む。前記サンプル中の前記ウイルスの存在の指標である。本明細書に記載のUSタイプ及びUSサブタイプE型肝炎ウイルスの発見に基づいて、サンプル中の前記ウイルスの存在を検出するために各種アッセイ、例えばタンパク質または核酸ベースのアッセイが作成され得る。タンパク質ベースのアッセイの例には慣用されているイムノアッセイが含まれ得、核酸ベースのアッセイの例には慣用されているプローブハイブリダイゼーションまたは核酸配列増幅アッセイが含まれ得、これらのアッセイはすべて当業界で公知であり、十分に検討されている。

【0044】

別の態様で、本発明は試薬、例えば抗体、ポリペプチド鎖を含むエピトープ、及び例えば予防目的または治療目的でUSタイプもしくはUSサブタイプE型肝炎ウイルスに対して患者を免疫化するためのワクチンを開発するために使用され得るヌクレオチド配列を提供する。

【0045】

I. 定義

本発明の理解を容易とするために、本明細書中で使用した用語は以下のように定義する。

【0046】

本明細書中、用語「USタイプ」E型肝炎ウイルスは、A型肝炎ウイルス(HAV)、B型肝炎ウイルス(HBV)、C型肝炎ウイルス(HCV)、D型肝炎ウイルス(HDV)及びG型肝炎ウイルス(HGV)とは血清学的に区別され、

少なくとも1個のオープンリーディングフレームを規定し、配列番号89の残基6307～6454により規定されるヌクレオチド配列と79.7%以上同一のヌクレオチド配列を有する一本鎖RNAゲノムを含むヒトウイルス（すなわち、ヒトを感染させ得る）を意味すると理解される。

【0047】

本明細書中、用語「USサブタイプ」E型肝炎ウイルスは、A型肝炎ウイルス（HAV）、B型肝炎ウイルス（HBV）、C型肝炎ウイルス（HCV）、D型肝炎ウイルス（HDV）及びG型肝炎ウイルス（HGV）とは血清学的に区別され、少なくとも1個のオープンリーディングフレームを規定し、配列番号89の残基6307～6454により規定されるヌクレオチド配列と90.5%以上同一のヌクレオチド配列を有する一本鎖RNAゲノムを含むヒトウイルス（すなわち、ヒトを感染させ得る）を意味すると理解される。

【0048】

本明細書中、用語「試験サンプル」は、検査すべきマーカー（例えば、抗体、抗原タンパク質またはペプチド、ヌクレオチド配列）を含む任意のサンプル、例えば生体サンプルを意味すると理解される。好ましい試験サンプルには、被験者から単離され得る組織または体液サンプルが含まれる。好ましい体液サンプルの例には、血液、血清、血漿、唾液、痰、血清、尿、便、胆汁、脊髄液、乳滲出液、腹水及び腹膜液が含まれる。別の好ましい試験サンプルは細胞系、より好ましくは哺乳動物細胞系である。最も好ましい細胞系はヒト胎児腎細胞系である。

【0049】

本明細書中、「オープンリーディングフレーム」または「ORF」は、1つ以上のポリペプチド鎖をコードし得るポリヌクレオチド配列の領域を意味すると理解される。この領域は、開始コドン（例えば、ATG（AUG））で始まり、終止コドン（例えば、TAA（UAA）、TAG（UAG）またはTGA（UGA））で終わる全コーディング配列またはその一部を表し得る。

【0050】

本明細書中、用語「ポリペプチド鎖」は、アミノ酸の分子鎖を意味すると理解され、特定長さの産物を指すのではない。よって、ポリペプチド鎖の定義にはペ

プチド、オリゴペプチド及びタンパク質が含まれる。

【0051】

本明細書中、用語「エピトープ」は「抗原決定基」と同義的に使用され、抗体可変領域により特異的に結合され得る（すなわち、約 10^5 M^{-1} 以上のアフィニティー、より好ましくは約 10^7 M^{-1} 以上のアフィニティーで結合する）抗原の少なくとも一部を意味すると理解される。多分、エピトープは該エピトープに独自の空間コンフォメーションで3個のアミノ酸を含み得る。一般的には、エピトープは少なくとも5個のアミノ酸を含み、より一般的には少なくとも8～10個のアミノ酸を含む。空間コンフォメーションを調べる方法は当業界で公知であり、その例にはX線結晶法及び二次元核磁気共鳴法が含まれる。

【0052】

ポリペプチドがポリペプチド鎖により規定される特定エピトープの抗体認識のために抗体に結合するとき、該ポリペプチドは抗体と「免疫学的に反応性」である。免疫学的反応性は、抗体結合（特に、抗体結合のカイネティック）により及び/または競合結合研究により調べられ得る。特定の抗体が第1抗原と免疫学的に反応性であるが第2の別の抗原とは免疫学的に反応しないか殆ど反応しない場合、2つの抗原は血清学的に区別されると考えられる。本明細書中、用語「アフィニティー」は、2つの分子間（例えば、抗体と抗原の間）の可逆的相互作用の尺度を意味すると理解される。アフィニティーが高いほど、2つの分子間の相互作用は強い。

【0053】

本明細書中、用語「検出可能部分」は、信号発生化合物、例えば色素原、触媒（例えば、酵素）、ルミネセンス化合物（例えば、ジオキセタン、アクリジニウム、フェナントリジニウム及びルミノール）、放射性元素及び肉眼で検出可能な標識を意味すると理解される。酵素の例には、アルカリホスファターゼ、ホースラディッシュペルオキシダーゼ、 β -ガラクトシダーゼ等が含まれる。特定の検出可能部分の選択は臨界的でないが、検出可能部分はそれ自体でまたは1つ以上の追加物質と組合せて信号を発生し得る。

【0054】

本明細書中、用語「固体支持体」は、プラスチック、誘導化プラスチック、磁性または非磁性金属、ガラスまたはシリコン表面を意味すると理解される。有用な表面の例には、試験管、微量滴定ウェル、シート、ビーズ、微粒子、チップ、ヒツジ（または他の適当な動物）赤血球または硬膜細胞(duracyte)が含まれる。適当な固体支持体は本発明の実施にとって臨界的でなく、当業者は選択可能である。固相上にペプチドを固定化するための好適な方法には、イオン相互作用、疎水性相互作用、共有相互作用等が含まれる。固体支持体は、捕捉試薬を引きつけ、固定化する固有の能力に関して選択され得る。或いは、固体支持体は捕捉試薬を引きつけ、固定化する能力を有する追加の受容体を有していてもよい。

【0055】

固体支持体は、検出抗体が接近し得るように十分な多孔度及び抗原を結合するために適当な表面アフィニティーを有する適当な多孔質材料からなり得ると考えられる。通常微孔性構造物が好ましいが、水和状態でゲル構造を有する材料も使用し得る。材料はすべてフィルム、シートまたはプラスチックのような適当な形態で使用され得、または紙、ガラス、プラスチックフィルムまたはファイバのような適当な不活性担体上にコートされたり、適当な不活性担体に結合またはラミネートされていてもよい。

【0056】

各種の他の固体支持体を用いる他の実施態様も本発明で考えられ、本発明の範囲に含まれる。例えば、高速液相免疫化学反応を実施するために、欧州特許出願公開第0 326 100号明細書及び同第0 406 473号明細書に記載されている負電荷を有するポリマーを用いて固定化可能な反応複合体を固定化するためのイオン捕捉方法を本発明では使用され得る。固定化可能な免疫複合体は、負電荷を有するポリアニオン/免疫複合体と予め処理した正電荷を有する多孔質マトリックスの間のイオン相互反応により反応混合物の残りに分離され、欧州特許出願公開第0 273 115号明細書に記載されている化学ルミネセント信号測定法に記載されているようなものを含めた公知の各種信号発生システムを用いて検出される。

【0057】

また、本発明の方法は、固相が磁性または非磁性微粒子からなる自動及び半自動システムを含めた微粒子方法を用いるシステムに使用するために改変され得る。前記システムには、1992年2月18日に付与された米国特許第5,089,424号明細書及び1993年9月14日に付与された米国特許第5,244,630号明細書に記載されているものが含まれる。

【0058】

イムノアッセイのための走査プローブ顕微鏡検査法の使用も本発明のモノクローナル抗体が容易に適用され得る技術である。走査プローブ顕微鏡検査法、特に原子力顕微鏡検査法では、捕捉相（例えば、本発明のモノクローナル抗体の少なくとも1つ）を固相に結合させ、走査プローブ顕微鏡を用いて固相表面上に存在するかもしれない抗原／抗体複合体を検出する。走査トンネル顕微鏡検査法を使用すると、多くのイムノアッセイシステムでは抗原／抗体複合体を検出するために通常使用しなければならない標識を使用しなくてすむ。特異的結合反応をモニターするためにSPMは多くの方法で使用され得る。1つの実施態様では、特異的結合パートナー（本発明のモノクローナル抗体であるアナライト特異的物質）の1つのメンバーを走査に適した表面に結合させる。アナライト特異的物質は、当業者に公知の方法に従ってプラスチックまたは金属表面の固相からなる試験片へ吸着させることにより結合され得る。或いは、特異的結合パートナー（アラナイト特異的物質）を誘導化プラスチック、金属、シリコンまたはガラスの固相からなる試験片に共有結合させてもよい。共有結合方法は当業者に公知であり、その中には特異的結合パートナーを試験片に非可逆的にリンクさせるための各種手段が含まれる。試験片がシリコンまたはガラスのときには、特異的結合パートナーを結合させる前に表面を活性化させなければならない。また、欧州特許出願公開第0322100号明細書及び同第0406473号明細書に記載されている技術及び化学を用いて特異的結合パートナーを試験片の表面上に固定化するためにポリ電解質相互作用を使用してもよい。共有結合による結合方法が好ましい。特異的結合メンバーを結合させた後、非特異的結合を最小限とするために表面を血清、タンパク質または他のブロッキング剤のような物質を用いて更に処理してもよい。表面をアッセイ目的の適合性を証明するために製造場所ま

たは使用時点で走査してもよい。走査方法は試験片の特異的結合特性を変更しないと予想される。

【0059】

本明細書中、用語「ヌクレオチド配列」または「核酸配列」は、任意の長さを有するヌクレオチドのポリマー形態を意味すると理解され、リボヌクレオチドでもデオキシヌクレオチドでもよい。この用語は分子の主構造を指す。よって、この用語には、二本鎖及び一本鎖DNA並びに二本鎖及び一本鎖RNAが含まれる。また、例えばメチル化及び/またはキャッピングによる修飾及びポリペプチドの未修飾形態も含まれる。

【0060】

本明細書中、用語「プライマー」は、標的ヌクレオチド配列にハイブリダイズし、DNAポリメラーゼ、RNAポリメラーゼまたは逆転写酵素により触媒されるヌクレオチド重合の開始点として機能し得る標的ヌクレオチド配列に相補性である特定のオリゴヌクレオチド配列を意味すると理解される。

【0061】

核酸断片を指すとき、前記核酸断片は検出の直線範囲内でHEV USタイプもしくはUSサブタイプのポリヌクレオチドまたはその変異体に特異的にハイブリダイズまたは特異的に結合すると見做され、このハイブリダイゼーションにより、同等量のHEV USタイプ、USサブタイプまたはその変異体以外に由来するポリヌクレオチドへのハイブリダイゼーションにより生ずるであろう信号に比してより強い信号が発生する。他に比してより強い信号は特定の検出方法により他の場合よりも測定可能である信号である。

【0062】

また、核酸断片を指すとき、前記核酸断片が(i)好ましいハイブリダイゼーション条件がより低い緊縮度を有するもの、より好ましくはより高い緊縮度を有するものであるManiatis, 第1版, p. 387 - 389 (1982)に記載されているような典型的なハイブリダイゼーション及び洗浄条件下、または(ii)標準のPCR条件(Saiki, R. K.ら)または“タッチダウン”PCR条件(Roux, K. H., Biotechniques, 16: 812 -

814(1994))下で特異的にハイブリダイズするならば、核酸断片は特定ハイブリダイゼーション条件下でハイブリダイズすると見做される。

【0063】

本明細書中、用語「プローブ」は、相補配列を有するサンプル中に存在する特定DNAまたはRNAを同定するために使用され得る配列を含有するヌクレオチドまたはヌクレオチドアナログ(例えば、PNA)を意味すると理解される。

【0064】

本明細書中、用語「PNA」は、標的の存在を調べるために本明細書に記載されているアッセイのような方法において使用され得るペプチド核酸アナログを意味すると使用されている。「MA」は、標的の存在を調べるために本明細書に記載されているアッセイのような方法において使用され得る「モルホリノアナログ」を指す。例えば、援用により本明細書に含まれるとする米国特許第5,378,841号明細書を参照されたい。PNAは通常RNA標的またはDNAに指向し得る中性に帯電した部分である。アッセイにおいて例えば本発明のDNAプローブの代わりにPNAプローブを使用しても、該DNAプローブを使用したときに得られる効果が達成され得ない。これらの効果には、製造性、大規模ラベリング、再現性、安定性、イオン強度の変化に対する非感受性、及びDNAまたはRNAを用いる方法で見られる酵素分解に対する耐性が含まれる。前記PNAをフルオレセイン、リボヌクレオチド、化学ルミセネント化合物等のような信号発生化合物で標識してもよい。従って、PNAまたは他の核酸アナログ(例えば、MA)はDNAまたはRNAの代わりにアッセイ方法で使用してもよい。本明細書ではDNAを用いるアッセイを説明しているが、アッセイ試薬において可能であったり所要により適切に変化させてRNAまたはDNAをPNAまたはMAで置換し得ることも日常的に行われていることである。

【0065】

核酸断片を指すとき、前記核酸断片は検出の直線範囲内でHEV USタイプもしくはUSサブタイプのポリヌクレオチドまたはその変異体に特異的にハイブリダイズまたは特異的に結合すると見做され、このハイブリダイゼーションにより、同等量のHEV USタイプ、USサブタイプまたはその変異体以外に由来

するポリヌクレオチドへのハイブリダイゼーションにより生ずるであろう信号に比してより強い信号が発生する。他に比してより強い信号は、特定の検出方法により他の場合よりも測定可能である信号である。

【0066】

また、核酸断片を指すとき、前記核酸断片が(i)好ましいハイブリダイゼーション条件がより低い緊縮度を有するもの、より好ましくはより高い緊縮度を有するものであるManiatis, 第1版, p. 387 - 389 (1982)に記載されているような典型的なハイブリダイゼーション及び洗浄条件下、または(ii)標準のPCR条件(Saiki, R. K.ら)または“タッチダウン”PCR条件(Roux, K. H., Biotechniques, 16: 812 - 814 (1994))下で特異的にハイブリダイズするならば、核酸断片は特定ハイブリダイゼーション条件下でハイブリダイズすると見做される。

【0067】

II. 検出方法及び試薬

本発明の方法は以下に詳記する各種のタンパク質ベースまたは核酸ベースのアッセイを使用し得ると考えられる。

【0068】

ウイルスまたはそのマーカーを検出するための試薬はUSタイプ及び/またはUSサブタイプE型肝炎ウイルス抗体、USタイプ及び/またはUSサブタイプ特異的ポリペプチド、またはUSタイプ及び/またはUSサブタイプE型肝炎ウイルスのゲノムの少なくとも一部を規定する核酸またはその相補的な核酸配列であり得ると考えられる。

【0069】

II - (i) タンパク質ベースのアッセイ

(a) マーカー抗体

ウイルスマーカーが当該患者の血流中を循環している抗-USタイプもしくは抗-USサブタイプ特異的抗体(例えば、IgGまたはIgM)であるならば、結合パートナーは前記マーカーに特異的に結合するエピトープを規定するポリペプチドであることが好ましいと考えられる。

【0070】

試験サンプル中の抗-USタイプもしくは抗-USサブタイプE型肝炎ウイルス抗体の存在を検出するための好ましいプロトコルでは、前記プロトコルは、好ましくは(a)少なくとも5個、より好ましくは少なくとも8個、更に好ましくは少なくとも15個、最も好ましくは少なくとも25個の隣接アミノ酸残基を含む免疫学的反応性のUSタイプもしくはUSサブタイプ特異的ポリペプチド鎖を含み、抗体に結合し得る抗原を用意するステップ、(b)抗体-抗原複合体を形成し得る条件で前記抗原を試験サンプルとインキュベートするステップ、及び(c)前記複合体の存在を検出するステップを含む。

【0071】

多くの各種USタイプもしくはUSサブタイプ特異的ポリペプチドが抗-USタイプもしくは抗-USサブタイプ抗体を検出するための結合パートナーとして有用であり得ると考えられる。例えば、前記ポリペプチド鎖は、配列番号91、92または93により規定されるアミノ酸配列、或いは配列番号91、92または93に示すポリペプチド鎖の好ましくは少なくとも5個、より好ましくは少なくとも8個、更に好ましくは少なくとも15個、最も好ましくは少なくとも25個の隣接アミノ酸残基を含み、ビルマ及びメキシコファミリーのメンバーの対応アミノ酸配列と比較してユニークなアミノ酸配列を表すその免疫学的に反応性の断片であり得ると考えられる。ビルマファミリー、すなわちビルマ様株は現在本明細書中でB1、B2、I1、I2、C1、C2、C3、C4及びP1と称する株を含み、メキシコファミリーは現在株M1を含む。

【0072】

結合パートナーは天然変異体を含めた配列番号91、92及び93により規定されるポリペプチドからなる群から選択されるポリペプチドであり得ると考えられる。本明細書中、配列番号91により規定されるポリペプチドに関して用語「天然変異体」は配列番号91の残基1～1698と少なくとも84%、好ましくは少なくとも86%、より好ましくは少なくとも89%、更に好ましくは少なくとも95%同一であるアミノ酸配列を意味すると解される。本明細書中、配列番号92により規定されるポリペプチドに関して用語「天然変異体」は配列番号9

2の残基1～660と少なくとも93%、好ましくは少なくとも95%、より好ましくは少なくとも98%同一であるアミノ酸配列を意味すると解される。本明細書中、配列番号93により規定されるポリペプチドに関して用語「天然変異体」は配列番号93の残基1～122と少なくとも85.4%、好ましくは少なくとも87.4%、より好ましくは少なくとも90.4%、更に好ましくは少なくとも95%同一であるアミノ酸配列を意味すると解される。

【0073】

更に、結合パートナーはORF1配列の一部によりコードされるポリペプチドであり得ると考えられる。ORF1配列によりコードされるタンパク質の例には、メチルトランスフェラーゼタンパク質、プロテアーゼタンパク質、Yドメインタンパク質、Xドメインタンパク質、ヘリカーゼタンパク質、超可変領域タンパク質及びRNA依存性RNAポリメラーゼタンパク質が含まれる。従って、有用なメチルトランスフェラーゼタンパク質は配列番号91の残基1～231に対して好ましくは少なくとも92.3%、より好ましくは少なくとも94.3%、最も好ましくは少なくとも97.3%の同一性を有すると考えられる。また、有用なプロテアーゼタンパク質は配列番号91の残基426～697に対して好ましくは少なくとも70.3%、より好ましくは少なくとも72.3%、最も好ましくは少なくとも75.3%の同一性を有すると考えられる。また、有用なYドメインタンパク質は配列番号91の残基207～424に対して好ましくは少なくとも94.6%、より好ましくは少なくとも96.6%、最も好ましくは少なくとも99.6%の同一性を有すると考えられる。また、有用なXドメインタンパク質は配列番号91の残基789～947に対して好ましくは少なくとも83.4%、より好ましくは少なくとも85.4%、最も好ましくは少なくとも88.4%の同一性を有すると考えられる。また、有用なヘリカーゼタンパク質は配列番号91の残基965～1197に対して好ましくは少なくとも92%、より好ましくは少なくとも94%、最も好ましくは少なくとも93%の同一性を有すると考えられる。また、有用な超可変領域タンパク質は配列番号91の残基698～788に対して好ましくは少なくとも28.7%、より好ましくは少なくとも30.7%、最も好ましくは少なくとも33.7%の同一性を有すると考えられ

る。また、有用なRNA依存性RNAポリメラーゼタンパク質は配列番号91の残基1212～1698に対して好ましくは少なくとも88.8%、より好ましくは少なくとも90.8%、最も好ましくは少なくとも93.8%の同一性を有すると考えられる。

【0074】

更に、結合パートナーは配列番号166、167または168により規定されるアミノ酸配列を有するポリペプチド鎖、或いは配列番号166、167または168に示すポリペプチド鎖の5個、好ましくは少なくとも8個、より好ましくは少なくとも15個、最も好ましくは少なくとも25個の隣接アミノ酸残基を含み、ビルマ及びメキシコファミリーのメンバーの対応アミノ酸配列と比較してユニークなアミノ酸配列を表すその免疫学的に反応性の断片であり得ると考えられる。同様に、結合パートナーは天然変異体を含めた配列番号166、167及び168からなる群から選択されるポリペプチドであり得ると考えられる。本明細書中、配列番号166により規定されるポリペプチドに関して用語「天然変異体」は配列番号166の残基1～1708と少なくとも83.9%、好ましくは少なくとも85.9%、より好ましくは少なくとも88.9%、最も好ましくは少なくとも95%同一であるアミノ酸配列を意味すると解される。本明細書中、配列番号167により規定されるポリペプチドに関して用語「天然変異体」は配列番号167の残基1～660と少なくとも93%、好ましくは少なくとも95%、最も好ましくは少なくとも98%同一であるアミノ酸配列を意味すると解される。本明細書中、配列番号168により規定されるポリペプチドに関して用語「天然変異体」は配列番号168の残基1～122と少なくとも85.4%、好ましくは少なくとも87.4%、より好ましくは少なくとも90.4%、更に好ましくは少なくとも95%同一であるアミノ酸配列を意味すると解される。

【0075】

更に、結合パートナーは、例えばメチルトランスフェラーゼタンパク質、プロテアーゼタンパク質、Yドメインタンパク質、Xドメインタンパク質、ヘリカーゼタンパク質、超可変領域タンパク質及びRNA依存性RNAポリメラーゼタンパク質を含めたHEV US-2 ORF1の一部によりコードされるポリペプ

チドまたはその変異体であり得ると考えられる。従って、有用なメチルトランスフェラーゼタンパク質は配列番号166の残基1～240に対して好ましくは少なくとも92.7%、より好ましくは少なくとも94.7%、最も好ましくは少なくとも97.7%の同一性を有すると考えられる。また、有用なプロテアーゼタンパク質は配列番号166の残基433～706に対して好ましくは少なくとも69.6%、より好ましくは少なくとも71.6%、最も好ましくは少なくとも74.6%の同一性を有すると考えられる。また、有用なYドメインタンパク質は配列番号166の残基216～433に対して好ましくは少なくとも94.6%、より好ましくは少なくとも96.6%、最も好ましくは少なくとも99.6%の同一性を有すると考えられる。また、有用なXドメインタンパク質は配列番号166の残基799～957に対して好ましくは少なくとも82.8%、より好ましくは少なくとも84.8%、最も好ましくは少なくとも87.8%の同一性を有すると考えられる。また、有用なヘリカーゼタンパク質は配列番号166の残基975～1207に対して好ましくは少なくとも92.8%、より好ましくは少なくとも94.8%、最も好ましくは少なくとも97.8%の同一性を有すると考えられる。また、有用な超可変領域タンパク質は配列番号166の残基707～798に対して好ましくは少なくとも27%、より好ましくは少なくとも29%、最も好ましくは少なくとも31%の同一性を有すると考えられる。また、有用なRNA依存性RNAポリメラーゼタンパク質は配列番号166の残基1222～1708に対して好ましくは少なくとも88.7%、より好ましくは少なくとも90.7%、最も好ましくは少なくとも93.7%の同一性を有すると考えられる。

【0076】

USタイプまたはUSサブタイプ特異的エピトープの同定に関し、USタイプもしくはUSサブタイプE型肝炎ウイルスのゲノムの少なくとも一部を規定する核酸配列及び/または該ゲノムの少なくとも一部によりコードされるアミノ酸配列を求めている当業者は当業界で公知であり、十分に検討されている慣用方法を用いて潜在的なエピトープをマップすることができると考えられる。所与の配列中の潜在的なエピトープを同定する市販のソフトウェアパッケージを用いる以外

に、前記ゲノムによりコードされるアミノ酸配列を抗原サイトが既に解明されているH E Vの他の株のゲノムによりコードされる配列と比較することにより潜在的なエピトープを同定することが可能である。例えば、米国特許第5,686,239号明細書、同第5,741,490号明細書及び同第5,770,689号明細書を参照されたい。現在同定されているエピトープを図1に示し、ここには当業界で8-5(配列番号93及び168)、4-2(配列番号93及び168の位置90~122)、SG3(配列番号175及び176)、3-2(配列番号92及び167の位置613~654)及び3-2e(配列番号92及び167の位置613~660)と称するエピトープが含まれている。抗原指数の計算方法はJameson及びWolf(CABIOS, 4(1), 181-186(1988))に記載されている。

【0077】

例えば、それぞれE型肝炎ゲノムのORF2及びORF3の一部によりコードされる興味深い2つのエピトープを以下に詳記し、3-2e及び4-2と称する。これらのエピトープは、H E Vのビルマ株(以下、B3-2e(配列番号172)及びB4-2(配列番号171)と称する)及びH E Vのメキシコ株(以下、M3-2e(配列番号170)及びM4-2(配列番号169)と称する)で同定された。アミノ酸配列比較に基づいて類似のエピトープがH E V US-1で同定され、これらを以下U3-2e(配列番号174)及びU4-2(配列番号173)と称する。同様にアミノ酸配列比較に基づいて類似のエピトープがH E V US-2で同定され、これらを以下US-2 3-3e(配列番号223)及びUS-2 4-2(配列番号224)と呼ぶ。

【0078】

加えて、潜在的なエピトープは当業界で公知であり、文献に詳しく記載されているスクリーニング方法を用いて同定され得る。例えば、H E V US-1またはH E V US-2ゲノムの全部または一部を規定する核酸配列に基づいて、発現後エピトープを同定すべくスクリーニングにかけられ得る発現ライブラリーを作成することが可能である。例えば、H E V US-1またはH E V US-2ゲノムの代表的な核酸断片をラムダ-gt11発現ベクターにクローニングして

、ラムダ - g t 1 1 ライブラリー（例えば、c D N A ライブラリー）を作成し得る。次いで、このライブラリーを、H E V U S - 1 または H E V U S - 2 に感染していると同定された患者から得た血清に特異的に結合し得るコード化エピトープについてスクリーニングする。例えば、G l o v e r , 「DNA クローニング方法 実際的アプローチ (DNA Cloning Techniques, A Practical Approach)」, p . 4 9 - 7 8 , I R L P r e s s (1 9 9 5 年) 発行を参照されたい。通常、約 $10^6 \sim 10^7$ ファージをスクリーニングし、これから陽性ファージを同定し、精製し、次いで H E V U S - 1 または H E V U S - 2 に感染していると既に同定された患者から得た血清への結合の特異性について試験する。患者からの血清または血漿中に存在する抗体に選択的に結合するファージを別の特性評価のために選択する。一旦同定されたら、当該アミノ酸配列を当業界で公知であり、文献に詳しく記載されている一般的な組み換え D N A 方法または一般的なペプチド合成方法を用いて大規模に産生させることができる。

【0079】

(b) マーカーポリペプチド

マーカーが U S タイプもしくは U S サブタイプウイルスまたはその特異的ポリペプチドであるならば、本発明を実施するのに有用な結合パートナーは前記ウイルスまたはマーカーポリペプチド上のエピトープに結合する抗体（例えば、ポリクローナルまたはモノクローナル抗体）であることが好ましいと考えられる。前記結合パートナーを検出可能部分で標識しても固体支持体上に固定化してもよい。特に、この実施態様を実施する際に有用な抗体は、ビルマ及びメキシコファミリーのメンバー中に存在する対応アミノ酸配列に比してユニークな U S タイプもしくは U S サブタイプ特異的ポリペプチド鎖に特異的に結合し得、前記ポリペプチド鎖は好ましくは少なくとも 5 個、より好ましくは 8 個、更に好ましくは少なくとも 1 5 個、最も好ましくは少なくとも 2 5 個の隣接アミノ酸残基の長さを有する。

【0080】

本発明の上記実施態様を実施する際に有用な抗体は天然変異体を含めた配列番号 9 1、9 2 及び 9 3 からなる群から選択されるポリペプチド鎖に特異的に結合

し得、ビルマ及びメキシコファミリーのメンバーの対応配列に比して前記ポリペプチド鎖に対してより高い結合アフィニティーを有することが好ましい。本発明を実施する際に有用な抗体は好ましくは配列番号173または175に示すアミノ酸配列を含むポリペプチド鎖に特異的に結合し得ると考えられる。前記抗体は更に、同一条件下で配列番号169または171に示すアミノ酸配列または配列番号175に対応するビルマ及びメキシコ株中の領域に対して低いアフィニティーを有し、最も好ましく結合しないとして特徴づけられる。また、本発明を実施する際に有用な抗体は好ましくは配列番号174または176に示すアミノ酸配列を含むポリペプチド鎖に特異的に結合し得ると考えられる。前記抗体は更に、同一条件下で配列番号170または172に示すアミノ酸配列または配列番号176に対応するビルマ及びメキシコ株中の領域に対して低いアフィニティーを有し、最も好ましく結合しないとして特徴づけられる。

【0081】

また、上記実施態様を実施する際に有用な抗体は天然変異体を含めた配列番号166、177及び168からなる群から選択されるポリペプチド鎖に特異的に結合し得、ビルマ及びメキシコファミリーのメンバーの対応配列に比して前記ポリペプチド鎖に対してより高い結合アフィニティーを有することが好ましいと考えられる。本発明を実施する際に有用な抗体は好ましくは配列番号223に示すアミノ酸配列を含むポリペプチド鎖に特異的に結合し得ると考えられる。前記抗体は更に、同一条件下で配列番号170または172に示すアミノ酸配列に対して低いアフィニティーを有し、最も好ましく結合しないとして特徴づけられる。また、本発明を実施する際に有用な抗体は好ましくは配列番号224に示すアミノ酸配列を含むポリペプチド鎖に特異的に結合し得ると考えられる。前記抗体は更に、同一条件下で配列番号169または171に示すアミノ酸配列に対して低いアフィニティーを有し、最も好ましく結合しないとして特徴づけられる。

【0082】

本明細書に記載の抗体またはその抗原結合断片は個別にUSタイプもしくはUSサブタイプ特異的抗原を検出するために提供され得る。本発明で提供される抗体（及びその抗原結合断片）の組合せも、USタイプもしくはUSサブタイプ特

異的抗原を分離するために異なる結合特異性を有する少なくとも2つの抗体を含む混合物、すなわちカクテル中の成分として一緒に使用してもよい。

【0083】

(c) 抗体作成

USタイプもしくはUSサブタイプE型肝炎ウイルスのORF 1、ORF 2及び/またはORF 3の少なくとも一部を規定するヌクレオチド配列または該ORF 1、ORF 2及び/またはORF 3の少なくとも一部によりコードされるアミノ酸配列を求めている当業者は、当業界で公知であり、文献に詳しく記載されている方法を用いて特異的抗体を作成することができると考えられる。例えば、*Practical Immunology*, Butt, N. R. 編, ニューヨークに所在のMarcel Dekker (1984年) 発行を参照されたい。簡単に説明すると、単離した標的タンパク質を使用して異種宿主(例えば、マウス、ブタ、ヤギまたは他の好適な哺乳動物)において抗体を産生する。好ましい抗体は、標的タンパク質上のエピトープに対して特異的に結合し、好ましくは該エピトープに対して 10^5 M^{-1} 以上、最も好ましくは 10^7 M^{-1} 以上の結合アフィニティーを有する抗体である。典型的には、標的タンパク質を、宿主中での抗体産生を強化し得る適当なアジュバントと組合わせ、宿主に対して例えば腹腔内投与により注射する。宿主免疫応答を刺激するのに適したアジュバントを使用するのが有利であり得る。通常使用されるアジュバントはフロイント完全アジュバント(例えば、カリフォルニア州サンジェゴに所在のCalbiochem Corp. またはニューヨーク州グランドアイランドに所在のGibcoから市販されている死滅乾燥させた微生物細胞を含むエマルジョン)である。複数回抗原注射が望ましいときには、その後の注射に抗原と不完全アジュバント(例えば、無細胞エマルジョン)の組合せを用いる。

【0084】

ポリクローナル抗体は、当該タンパク質に対する抗体を含む血清を抽出することにより抗体産生宿主から単離され得る。モノクローナル抗体は、所望抗体を産生する宿主細胞を単離し、前記細胞をミエローマ細胞と免疫分野で公知の一般的な方法(例えば、Kohler及びMilstein, *Nature*, 256:

495(1975)参照)を用いて融合し、標的タンパク質と特異的に反応し、所望結合アフィニティーを有するハイブリッド細胞(ハイブリドーマ)についてスクリーニングすることにより作成され得る。

【0085】

加えて、小ペプチドを使用するとき、その免疫原性は固体支持体にカップリングさせることにより強化され得ると考えられる。例えば、ポリペプチドのエピトープもしくは抗原領域または断片は比較的小さく、約8~10アミノ酸以下の長さを有し得る。3アミノ酸ほどの小さい断片が抗原領域を特徴づけ得る。前記ポリペプチドは、当該ポリペプチドが折り畳まれると適正なエピトープを与えるが抗原であるには非常に小さいときにはポリペプチドは好適な担体分子に結合され得る。

【0086】

前記使用のための好ましい結合試薬及び方法は当業者で公知であり、それにはN-スクシンイミジル-3-(2-ピリジルチオ)プロピオネート(SPD P)及びスクシンイミジル4-(N-マレイミドメチル)シロヘキサン-1-カルボキシレート(SMCC)が含まれるが、これらに限定されない。更に、スルフヒドリル基を結合するポリペプチドはシステイン残基を添加することにより修飾され得る。前記試薬により、試薬と1つのタンパク質上のペプチドシステイン残基の間にジスルフィド結合及びリシン上のイプシロンアミノを介してアミド結合が形成され、または他の遊離アミノ基が形成される。いろいろなジスルフィド/アミド形成剤が公知である。他の二官能性カップリング剤はジスルフィド結合よりもむしろチオエステルを形成する。多くのチオエーテル形成剤が市販されており、当業者に公知である。カルボキシル基は、スクシンイミドまたは1-ヒドロキシ-2-ニトロ-4-スルホン酸ナトリウムと混合することにより活性化され得る。それ自体宿主にとって有害な抗体を産生しない担体も使用可能である。好適な担体には、タンパク質、ポリサッカライド(例えば、ラテックス官能化セファロース)、アガロース、セルロース、セルロースビーズ、高分子アミノ酸(例えば、ポリグルタミン酸)、ポリリシン及び非酸コポリマー及び不活性ウイルス粒子が含まれる。タンパク質基質の例には、血清アルブミン、キーホールリンペッ

トヘモシアニン、免疫グロブリン分子、チログロブリン、オボアルブミン、破傷風トキソイド及び当業者に公知の別のタンパク質が含まれる。

【0087】

加えて、好ましいエピトープに対する結合アフィニティーを高めるために結合ドメインのアミノ酸配列が操作されている生合成した抗体結合ドメインも本発明を実施するのに有用であり得ると考えられる。この製造の詳細説明は、例えば *Practical Immunology*, Butt, W. R. 編, ニューヨークに所在の Marcel Dekker (1984年) 発行に見つけられ得る。場合により、一価抗体断片 (例えば、Fab または Fab' 断片) が利用され得る。更に、例えば米国特許第 5,091,513 号明細書及び同第 5,132,405 号明細書に記載されているように、1) 非共役結合またはジスルフィド結合した合成 V_H 及び V_L ダイマー、2) 共役結合した $V_H - V_L$ 一本鎖抗体結合サイト、3) 個別の V_H または V_L ドメイン、または 4) 一本鎖結合サイトを含む遺伝子工学処理した生合成抗体結合サイトも利用され得る。

【0088】

US タイプもしくは US サブタイプ E 型肝炎ウイルスエピトープに結合する無傷の抗体 (例えば、モノクローナルまたはポリクローナル抗体)、抗体断片または生合成抗体結合サイトは診断及び予防分野で有用であり、受動免疫治療においても有用である。

【0089】

(d) アッセイフォーマット

抗-US タイプもしくは抗-US サブタイプ E 型肝炎ウイルス特異的抗体または US タイプもしくは US サブタイプ E 型肝炎特異的エピトープに対する抗体を含む血清と免疫学的に反応するポリペプチドが当該試験サンプル中の前記ウイルスの存在を検出するためのイムノアッセイにおいて有用であると考えられる。更に、サンプル中の US タイプもしくは US サブタイプ E 型肝炎ウイルスの存在は当業界で公知であり、文献に詳しく記載されている広範囲のイムノアッセイ方法、例えば直接アッセイ、サンドイッチアッセイ及び/または競合アッセイを用いて検出され得ると考えられる。各種の好ましいアッセイフォーマットを以下に詳

記する。

【0090】

1つの好ましいフォーマットでは、アッセイはサンドイッチフォーマットを使用する。サンドイッチイムノアッセイは、良好な検出限界を有する標識が使用されるならば通常高特異性且つ非常に感受性である。免疫学的アッセイ設計、理論及びプロトコルの詳しい説明は、*Practical Immunology*, Butt, W. R. 編, ニューヨークに所在の Marcelle Deller (1984年) 発行を含めた多数の文献に見つけることができる。

【0091】

1つのタイプのサンドイッチフォーマットでは、固体支持体上に固定化されており、抗-USタイプもしくは抗-USサブタイプE型肝炎ウイルス抗体(マーカー)と免疫学的に反応するポリペプチド(結合パートナー)をUSタイプもしくはUSサブタイプE型肝炎ウイルスに感染していると疑われる患者からの試験サンプルと接触させて混合物を形成する。次いで、この混合物をポリペプチド/抗体複合体を形成するのに十分な時間及び条件でインキュベートする。次いで、試験サンプル抗体に特異的に結合するモノクローナルまたはポリクローナル抗体またはその断片を含み、検出可能な部分で標識したインジケータ試薬を前記の抗原/抗体複合体と接触させて、第2混合物を形成する。次いで、この第2混合物を抗原/抗体/抗体複合体を形成するのに十分な時間及び条件でインキュベートする。試験サンプル中に存在するかもしれない抗-USタイプもしくは抗-USサブタイプE型肝炎ウイルス抗体の存在は、固体支持体に固定化した検出可能な部分の存在を検出することにより調べられる。試験サンプル中に存在する抗体の量は発生する信号に比例する。ビオチンとアンチビオチン、ビオチンとアビジン、ビオチンとストレプトアビジン等を使用すると、本明細書に記載のアッセイにおいて発生した信号が強化され得る。

【0092】

上記アッセイの別のフォーマットでは、免疫学的に反応性のポリペプチドを固体支持体に間接的に、すなわち該ポリペプチドに特異的に結合するモノクローナルまたはポリクローナル抗体またはその断片を介して固定化し得る。或いは、別

のフォーマットでは、試験サンプル抗体、すなわちマーカー抗体（例えば、IgGまたはIgM）に特異的に結合し、固体支持体上に固定化されている抗体またはその抗原結合断片が抗体／抗体複合体を形成するのに十分な時間及び条件で試験サンプルと接触するようにアッセイ成分を逆配置で使用してもよい。次いで、インジケータ試薬、例えば捕捉された試験サンプル抗体と免疫学的に反応し且つ検出可能部分で標識したUSタイプもしくはUSサブタイプE型肝炎ポリペプチドを抗体／抗体／抗原複合体を形成し得るのに十分な時間及び条件で前記抗体／抗体複合体とインキュベートして、第2混合物を形成する。上記したように、固体支持体上に固定化されている捕捉抗体またはその抗原結合断片により捕捉される、試験サンプル中に存在するかもしれない抗体の存在は検出可能な部分から発生する測定可能な信号を検出することにより調べられる。

【0093】

上記したサンドイッチアッセイも、試験サンプル中のUSタイプもしくはUSサブタイプE型肝炎ウイルスまたはその免疫学的に反応性のポリペプチドの存在を上記アッセイフォーマットのルーチン改変により調べるために使用され得ると考えられる。前記改変は当業者に公知であると考えられる。

【0094】

上記のサンドイッチアッセイに加えて、本発明の実施において競合アッセイをも使用され得ると考えられる。このフォーマットでは、USタイプもしくはUSサブタイプE型肝炎特異的ポリペプチド鎖に特異的に結合する少なくとも2つの抗体（好ましくは、モノクローナル抗体）の1つまたはその組合せをUSタイプもしくはUSサブタイプ特異的タンパク質に対する抗体を検出するために競合プローブとして使用し得る。例えば、マーカーに対する結合パートナーとして作用する第1のHEV US-1特異的ポリペプチド鎖（例えば、本明細書に記載されているポリペプチドの1つ）を固体支持体上に固定化する。次いで、HEV US-1抗原に対する抗体を含んでいると疑われる試験サンプルを、固体支持体上に固定化された抗原／抗体複合体を形成するのに十分な時間及び条件で固体支持体及び例えば固定化されているHEV US-1特異的ポリペプチド鎖に結合し且つ検出可能部分で標識した単離抗-USタイプもしくは抗-USサブタイプ

抗体を含むインジケータ試薬とインキュベートする。マーカー固体が試験サンプル中に存在するならば、前記マーカー抗体は固定化ポリペプチドに結合するために標識インジケータ試薬と競合する。試験サンプル中に存在するマーカー抗体の量が増加すると、固定化ポリペプチドに結合する標識インジケータ試薬の量は減少する。固相に結合するインジケータ試薬の量の減少は定量化され得る。非A、非B、非C、非D、非E肝炎が陰性であると確認された試験サンプルから発生した信号と比較して信号が低下していると認められたならば、試験サンプル中の抗-H E V U S - 1抗体の存在が明らかとなる。試験サンプル中のU S タイプもしくはU S サブタイプクラスに入る他のE型肝炎ウイルスの存在を同定するために同様のプロトコルを使用してもよいと考えられる。

【0095】

更に別の検出方法では、本発明の抗体は固定組織切片及び固定細胞中のU S タイプもしくはU S サブタイプE型肝炎特異的抗原の存在を免疫組織化学的分析により検出するために使用され得る。抗体が検出可能部分（例えば、フルオレセイン、金コロイド、ホースラディッシュペルオキシダーゼ、アルカリホスファターゼ等）で直接標識されているかまたは例えば検出可能部分で標識した第2抗体を用いて間接的に標識されている細胞化学分析も本発明を実施する際に使用され得る。

【0096】

別のアッセイフォーマットでは、抗体及び/または抗原の存在は例えば欧州特許出願公開第0 4 7 3 0 6 5号明細書に記載されているような同時アッセイを用いて検出され得る。例えば、試験サンプルを、(i) 固体支持体上に固定化した第1アナライトに対して特異的な第1結合メンバーを含む第1アナライトに対する捕捉試薬及び(i i) 第2の別の固体支持体上に固定化した第2アナライトに対する第1結合メンバーを含む第2アナライトに対する捕捉試薬と同時に接触させて混合物を生成する。次いで、この混合物を捕捉試薬/第1アナライト複合体及び捕捉試薬/第2アナライト複合体を形成するのに十分な時間及び条件でインキュベートする。こうして形成した複合体を、検出可能部分で標識した第1アナライトに対して特異的な結合対のメンバーを含む第1インジケータ試薬及び

検出可能部分で標識した第2アナライトに対して特異的な結合対のメンバーを含む第2インジケータ試薬と接触させて第2混合物を生成する。次いで、この第2混合物を捕捉試薬/第1アナライト/第1インジケータ試薬複合体及び捕捉試薬/第2アナライト/第2インジケータ試薬複合体を形成するのに十分な時間及び条件でインキュベートする。1つ以上のアナライトの存在を、固相の一方または両方上に形成された複合体により発生する信号を試験サンプル中に存在する1つ以上のアナライトの存在の指標として検出することにより調べる。

【0097】

他のアッセイシステムでは、特定抗体とウイルスタンパク質（例えば、ペプチド等）の接触によりウイルスゲノム（または、その断片）を包囲するUSタイプもしくはUSサブタイプE型肝炎ウイルス粒子またはサブウイルス粒子に特異的に結合する抗体が使用され得る。次いで、捕捉された粒子は、ウイルスゲノムが試験サンプル中に存在するかを調べるためにLCRまたはPCRのような方法により分析され得る。抗原捕捉増幅法を用いる利点は、ウイルスゲノムを試験サンプル中の他の分子から特異的抗体を用いて分離し得ることである。前記方法は1995年9月20日に公開された欧州特許出願公開第0 672 176号明細書に記載されている。

【0098】

一般的に、イムノアッセイ設計には、非特異的相互作用の産物とは容易に区別され得る複合体を形成するように標的タンパク質に対して十分高い結合特異性を有する抗体（例えば、モノクローナルまたはポリクローナル抗体、またはその抗原結合断片）の作成が含まれる。通常、抗体結合特異性が高ければ、より低濃度の標的が検出され得る。

【0099】

本発明のポリペプチド及び抗体試薬は、USタイプもしくはUSサブタイプE型肝炎ウイルスに結合する抗原形態または抗体の存在を検出するための本明細書に記載のアッセイを実施するために使用され得る。イムノアッセイでの使用に加えて、上記ポリペプチドは単独で、または実験動物において抗体を産生するために使用されるアジュバンドと組み合わせて、または患者を予防または治療目的で

免疫化するためのワクチンとして医薬的に許容され得る担体と組み合わせて使用され得ると考えられる。また、イムノアッセイでの使用に加えて、本発明の抗体は例えば患者を治療または予防目的で受動免疫化するために使用される医薬的に許容され得る担体と組み合わせて使用され得ると考えられる。後者の使用は以下のセクション(III)に詳記する。本発明の抗体は、治療または他の類似の用途で使用されるキメラ抗体の作成のためにも使用可能である。

【0100】

適当な試薬を含有する免疫診断に好適なキットは、例えば当該特定エピトープを規定するポリペプチドまたは前記エピトープに結合する抗体を含めた適当な材料を好適な容器中に包装することにより構成され得る。更に、前記キットは場合により追加の試薬、例えば適当な検出システム及び緩衝液を含み得る。

【0101】

更に、前記抗体(好ましくは、モノクローナル抗体)をCNBr活性化セファロースに類似のマトリックスに結合させ、例えば組み換え及び天然のウイルス抗原及びタンパク質を精製するために細胞培養物または生体組織(例えば、血液及び肝臓)由来のUSタイプもしくはUSサブタイプE型肝炎タンパク質をアフィニティー精製するために使用してもよい。

【0102】

III - (ii) 核酸ベースのアッセイ

マーカーがUSタイプもしくはUSサブタイプ特異的ヌクレオチド配列であるならば、結合パートナーはマーカー配列または該配列に隣接する領域に特異的にハイブリダイズするヌクレオチド配列またはそのアナログであることが好ましいと考えられる。本明細書に記載されているユニークなポリヌクレオチド配列に基づいて、結合パートナーは、ビルマ及びメキシコファミリーの対応ヌクレオチド配列と比較してユニークなUSタイプもしくはUSサブタイプ特異的ヌクレオチド配列に相補的なヌクレオチド配列、例えばUSタイプもしくはUSサブタイプE型肝炎ウイルスのORF1配列、ORF2配列またはORF3配列の少なくとも一部に相補的なヌクレオチド配列またはそのアナログであり得ると考えられる。更に、E型肝炎のビルマ及びメキシコファミリーのゲノムに比してユニークな

USタイプ及びUSサブタイプE型肝炎ウイルスのゲノムの非コーディング部分が本発明の実施において有用なマーカーを提供すると考えられる。ヌクレオチド配列（プライマーまたはプローブ）はハイブリダイゼーション及び/または増幅によりUSタイプ及びUSサブタイプ特異的配列を検出できる長さを有し、当業界で公知であり、十分検討されている自動化オリゴヌクレオチド合成法を含めたルーチンな標準方法を用いて製造され得る。HEV US-1ゲノムのユニーク部分の補体が満足である。プローブとして使用するには完全相補性が望ましいが、断片の長さを長くするので完全相補性が不必要な場合がある。

【0103】

同様に、結合パートナーは、標的配列に特異的にハイブリダイズし得、好ましくは8～100ヌクレオチド、より好ましくは10～75ヌクレオチド、最も好ましくは15～50ヌクレオチドを含むポリヌクレオチド配列（例えば、DNA、RNAまたはPNA配列）であり得ると考えられる。標的配列とは、USタイプもしくはUSサブタイプE型肝炎ウイルスのゲノムの少なくとも一部を規定するヌクレオチド配列または該配列に相補的な配列であり得ると解される。ハイブリダイゼーション反応のために選択される特定の緊縮条件が結合パートナー核酸配列と標的配列の相補度、結合配列の組成及び結合配列の長さに大いに依存することは当業界で公知である。緊縮条件を決めるパラメーターは当業者に公知であるかまたは一般文献（例えば、Maniatisら、「分子クローニング：実験室マニュアル(Molecular Cloning: A Laboratory Manual)」, ニューヨークに所在のCold Spring Harbor Press (1989年) 発行参照）から容易に確認されると認められる。

【0104】

本発明で提供される配列は、試験サンプル中の核酸を検出するためのアッセイにおいて使用され得るプローブを作成するために使用され得る。前記プローブは当該ポリヌクレオチドの保存ヌクレオチド領域から、または当該ポリヌクレオチドの非保存ヌクレオチド領域から設計され得る。アッセイにおいて最適なプローブの設計は日常作業者の技量の範囲内である。通常、核酸プローブは最大特異性を所望するときには非保存またはユニーク領域から作成され、核酸プローブは例

えば多遺伝子ファミリーの各種メンバーまたはマウスやヒトのような関連種に密接に関連するヌクレオチド領域をアッセイする場合には保存領域から作成される。

【0105】

1つの好ましいプロトコルは、試験サンプル中のUSタイプもしくはUSサブタイプE型肝炎ウイルスの存在または非存在の検出方法を提供する。この方法は、(a)USタイプもしくはUSサブタイプ単離物由来の少なくとも15隣接ヌクレオチドを含み、E型肝炎ビルマ及びメキシコファミリーの他のメンバー中には存在しないポリヌクレオチド配列を含むプローブを用意するステップ、(b)前記プローブと試験サンプル中に存在する非USタイプ及び非USサブタイプ肝炎ポリヌクレオチド配列との間で実質的にポリヌクレオチド二重らせんを形成せずに、プローブとその補体との間でポリヌクレオチド二重らせんを形成し得る条件下で試験サンプルをプローブと接触させるステップ、及び(c)前記プローブを含むポリヌクレオチド二重らせんの存在を検出するステップを含む。

【0106】

好ましいヌクレオチド配列は、配列番号89のヌクレオチド残基1～5097またはその天然に存在する配列変異体を含み得る。この配列に関して、用語「天然に存在する配列変異体」には配列番号89の残基1～5097と少なくとも73.3%、好ましくは少なくとも75.3%、より好ましくは少なくとも78.3%、最も好ましくは少なくとも95%同一の核酸配列が含まれる。他の好ましいマーカ―または結合パートナー配列は、配列番号89のヌクレオチド残基5132～7114またはその天然に存在する配列変異体を含み得る。この配列に関して、用語「天然に存在する配列変異体」には配列番号89の残基5132～7114と少なくとも87.4%、好ましくは少なくとも89.4%、より好ましくは少なくとも92.4%、最も好ましくは少なくとも95%同一の核酸配列が含まれる。他の好ましいマーカ―または結合パートナー配列は、配列番号89のヌクレオチド残基5094～5462またはその天然に存在する配列変異体を含み得る。この配列に関して、用語「天然に存在する配列変異体」には配列番号89の残基5094～5462と少なくとも88.3%、好ましくは少なくとも9

0.3%、より好ましくは少なくとも93.3%、最も好ましくは少なくとも95%同一の核酸配列が含まれる。

【0107】

更に、有用なヌクレオチド配列には、例えばメチルトランスフェラーゼタンパク質、プロテアーゼタンパク質、Yドメインタンパク質、Xドメインタンパク質、ヘリカーゼタンパク質、超可変領域タンパク質及びRNA依存性RNAポリメラーゼタンパク質からなる群から選択されるタンパク質またはその変異体をコードするORF1配列の一部が含まれる得ると考えられる。従って、ORF1の有用なメチルトランスフェラーゼコーディング領域は配列番号89の残基1~693に対して好ましくは少なくとも78%、より好ましくは少なくとも80%、最も好ましくは少なくとも83%の同一性を有すると考えられる。また、ORF1の有用なプロテアーゼコーディング領域は配列番号89の残基1270~2091に対して好ましくは少なくとも66.1%、より好ましくは少なくとも68.1%、最も好ましくは少なくとも71.1%の同一性を有すると考えられる。また、ORF1の有用なYドメインコーディング領域は配列番号89の残基619~1272に対して好ましくは少なくとも80%、より好ましくは少なくとも82%、最も好ましくは少なくとも85%の同一性を有すると考えられる。また、ORF1の有用なXドメインコーディング領域は配列番号89の残基2365~2841に対して好ましくは少なくとも73.5%、より好ましくは少なくとも75.5%、最も好ましくは少なくとも78.5%の同一性を有すると考えられる。また、ORF1の有用なヘリカーゼコーディング領域は配列番号89の残基2893~3591に対して好ましくは少なくとも77.5%、より好ましくは少なくとも79.5%、最も好ましくは少なくとも81.5%の同一性を有すると考えられる。また、ORF1の有用な超可変領域コーディング領域は配列番号89の残基2092~2364に対して好ましくは少なくとも51.2%、より好ましくは少なくとも53.2%、最も好ましくは少なくとも56.2%の同一性を有すると考えられる。また、ORF1の有用なRNA依存性RNAポリメラーゼコーディング領域は配列番号89の残基3634~5094に対して好ましくは少なくとも76.3%、より好ましくは少なくとも78.3%、最も好まし

くは少なくとも81.3%の同一性を有すると考えられる。

【0108】

好ましいヌクレオチド配列は、配列番号164のヌクレオチド残基36～5162またはその天然に存在する配列変異体を含み得る。この配列に関して、用語「天然に存在する配列変異体」には配列番号164の残基36～5162と少なくとも73.6%、好ましくは少なくとも75.6%、より好ましくは少なくとも78.6%、最も好ましくは少なくとも95%同一の核酸配列が含まれる。他の好ましいマーカ―または結合パートナー配列は、配列番号164のヌクレオチド残基5197～7179またはその天然に存在する配列変異体を含み得る。この配列に関して、用語「天然に存在する配列変異体」には配列番号164の残基5197～7179と少なくとも80.7%、好ましくは少なくとも82.7%、より好ましくは少なくとも85.7%、最も好ましくは少なくとも95%同一の核酸配列が含まれる。他の好ましいマーカ―または結合パートナー配列は、配列番号164のヌクレオチド残基5159～5527またはその天然に存在する配列変異体を含み得る。この配列に関して、用語「天然に存在する配列変異体」には配列番号164の残基5159～5527と少なくとも87.9%、好ましくは少なくとも89.9%、より好ましくは少なくとも92.9%、最も好ましくは少なくとも95%同一の核酸配列が含まれる。

【0109】

更に、有用なHEV US-2ヌクレオチド配列には、例えばメチルトランスフェラーゼタンパク質、プロテアーゼタンパク質、Yドメインタンパク質、Xドメインタンパク質、ヘリカーゼタンパク質、超可変領域タンパク質及びRNA依存性RNAポリメラーゼタンパク質からなる群から選択されるタンパク質またはその変異体の一部をコードするORF1配列の一部が含まれる得ると考えられる。従って、ORF1の有用なメチルトランスフェラーゼコーディング領域は配列番号164の残基36～755に対して好ましくは少なくとも79.5%、より好ましくは少なくとも81.5%、最も好ましくは少なくとも84.5%の同一性を有すると考えられる。また、ORF1の有用なプロテアーゼコーディング領域は配列番号164の残基1332～2153に対して好ましくは少なくとも6

6.1%、より好ましくは少なくとも68.1%、最も好ましくは少なくとも71.1%の同一性を有すると考えられる。また、ORF1の有用なYドメインコーディング領域は配列番号164の残基680~1334に対して好ましくは少なくとも80.7%、より好ましくは少なくとも82.7%、最も好ましくは少なくとも85.7%の同一性を有すると考えられる。また、ORF1の有用なXドメインコーディング領域は配列番号164の残基2430~2906に対して好ましくは少なくとも73.7%、より好ましくは少なくとも75.7%、最も好ましくは少なくとも78.7%の同一性を有すると考えられる。また、ORF1の有用なヘリカーゼコーディング領域は配列番号164の残基2958~3656に対して好ましくは少なくとも76.4%、より好ましくは少なくとも78.4%、最も好ましくは少なくとも81.4%の同一性を有すると考えられる。また、ORF1の有用な超可変領域コーディング領域は配列番号164の残基2154~2429に対して好ましくは少なくとも50.4%、より好ましくは少なくとも52.8%、最も好ましくは少なくとも55.8%の同一性を有すると考えられる。また、ORF1の有用なRNA依存性RNAポリメラーゼコーディング領域は配列番号89の残基3699~5159に対して好ましくは少なくとも76.8%、より好ましくは少なくとも78.8%、最も好ましくは少なくとも81.8%の同一性を有すると考えられる。

【0110】

他の有用なヌクレオチド配列は、配列番号93、168、173、174、175、176、223及び224からなる群から選択されるアミノ酸配列をコードするヌクレオチド配列及び該配列に相補的なヌクレオチド配列を含む。

【0111】

本発明で提供される核酸配列は、一般的な核酸ベースのアッセイ、例えば以下に詳記するポリメラーゼ連鎖反応(PCR)及び/またはプロットハイブリダイゼーション研究により試験サンプル中のUSタイプもしくはUSサブタイプE型肝炎ウイルスの存在を調べるために使用され得ると考えられる。核酸ベースのアッセイでの使用に加えて、上記核酸配列はその後当該宿主細胞(例えば、天然痘またはミコバクテリア)に形質転換またはトランスフェクトされ得るベクターに

統合され得ると考えられる。生じた宿主細胞は、その後医薬的に許容され得る担体と組合わされ、例えば哺乳動物を特定のUSタイプもしくはUSサブタイプE型肝炎ウイルスに対して予防または治療目的で免疫化するための組み換えワクチンとして使用される。

【0112】

ポリメラーゼ連鎖反応（PCR）は核酸またはその混合物中に含まれる所望の核酸配列（標的）を増幅させる方法である。PCRでは、プライマー対は通常標的核酸の相補鎖の外端でハイブリダイズするために過剰に使用される。プライマーはそれぞれ、鋳型として標的核酸を用いてポリメラーゼ（例えば、熱安定性ポリメラーゼ）により延長される。延長産物は元の配列鎖から解離後それ自体標的配列となる。その後、新しいプライマーをハイブリタイズし、ポリメラーゼにより延長するが、このサイクルは標的配列分子の数を等比的に増加させるために繰り返される。PCRは米国特許第4,683,195号明細書及び同第4,683,202号明細書に開示されている。

【0113】

リガーゼ連鎖反応（LCR）は核酸増幅の代替方法である。LCRでは、2つの一次（第1及び第2）プローブ及び2つの二次（第3及び第4）プローブを含むプローブ対を使用し、前記プローブをすべて標的核酸配列に対してモル過剰で使用する。第1プローブは標的鎖の第1セグメントにハイブリダイズし、第2プローブは標的鎖の第2セグメントにハイブリダイズするが、前記の第1及び第2セグメントは一次プローブが5'ホスフェート-3'-ヒドロキシ関係で相互に隣接するように且つリガーゼが2つのプローブを共役的に融合するかまたは連結して融合産物とし得るように隣接している。加えて、第3（二次）プローブは第1プローブの一部にハイブリダイズし得、第4（二次）プローブは同様に隣接して第2プローブの一部にハイブリダイズし得る。一旦一次プローブの連結鎖が標的鎖から分離したら、該連結鎖は第3及び第4プローブにハイブリダイズし、連結して相補的二次連結産物を形成する。連結産物は標的またはその補体と機能的に同等である。ハイブリダイゼーション及び連結のサイクルを繰り返すことにより、標的配列は増幅される。この技術は1989年6月16日に公開されたK.

B e c k m a n の欧州特許出願公開第320 308号明細書及び1991年7月31日に公開されたK . B e c k m a n の欧州特許出願公開第439 182号明細書により詳しく記載されている。

【0114】

mRNAを増幅させるために、mRNAをcDNAに逆転写した後ポリメラーゼ連鎖反応(RT-PCR)すること；米国特許第5,322,770号明細書に記載されているように両ステップに対して1つの酵素を使用すること；またはR . L . M a r s h a l l ら, P C R M e t h o d s a n d A p p l i c a t i o n s , 4 : 8 0 - 8 4 (1 9 9 4) に記載されているようにmRNAをcDNAに逆転写した後非対称ギャブリガーゼ連鎖反応(RT-AGLCR)を行うことも本発明の範囲内である。

【0115】

本発明において使用可能な他の公知の増幅方法には、P r o c . N a t l . A c a d . S c i . U S A , 8 7 : 1 8 7 4 - 1 8 7 8 (1 9 9 0) 及びN a t u r e , 3 5 0 (N o . 6 3 1 3) : 9 1 - 9 2 (1 9 9 1) に記載されている所謂“NASBA”または“3SR”；欧州特許出願公開第4544640号明細書に記載されているQ-ベーター増幅；G . T . W a l k e r ら, C l i n . C h e m . , 4 2 : 9 - 1 3 (1 9 9 6) 及び欧州特許出願公開第684315号明細書に記載されている鎖置換増幅；及び国際特許出願公開第9322461号パンフレットに記載されている標的媒介増幅が含まれるが、これらに限定されない。

【0116】

1つの実施態様で、本発明は包括的に、標的ポリヌクレオチド配列を含んでいると疑われる試験サンプルを増幅プライマーを含む増幅反応試薬及びアンプリコン配列の内部領域とハイブリダイズし得る検出プローブと接触させるステップを含む。本発明の方法に従って使用されるプローブ及びプライマーは、前記プローブが1つのタイプの標識で標識され、プライマーが他のタイプの標識で標識されるように捕捉及び検出標識で標識される。また、プライマー及びプローブは、プローブ配列がプライマー配列よりも低い融点を有するように選択される。増幅試

薬、検出試薬及び試験サンプルを増幅条件下に置き、こうすると標的配列の存在下で標的配列（アンプリコン）のコピーが作成される。その後、二本鎖アンプリコンを熱変性すると、一本鎖アンプリコンメンバーが形成される。一本鎖アンプリコンメンバーが形成されたら、混合物を冷却して、プローブと一本鎖アンプリコンメンバーの間で複合体を形成する。

【0117】

プローブ/一本鎖アンプリコンメンバーハイブリッドを形成した後、前記ハイブリッドを検出する。標準の異種アッセイフォーマットがプライマー及びプローブ上に存在する検出標識及び捕捉標識を用いてハイブリッドを検出するのに適している。前記ハイブリッドは捕捉標識により固相試薬に結合され得、検出標識により検出され得る。検出標識が直接検出可能な場合には、固相上のハイブリッドの存在は、所要により標識により検出可能な信号を生じさせ、その信号を検出することにより検出され得る。標識が直接検出可能でない場合には、捕捉されたハイブリッドを通常直接検出可能な標識に結合した結合メンバーを含むコンジュゲートと接触させ得る。前記コンジュゲートは複合体と結合するようになり、複合体上に存在するコンジュゲートは直接検出可能な標識で検出され得る。よって、固相試薬上のハイブリッドの存在が調べられ得る。当業者は、ハイブリダイズされていないアンプリコンまたはプローブ及び非結合コンジュゲートを除去するために洗浄ステップを使用し得ることを認識している。

【0118】

標的配列を検出するための試験サンプルは当業者に公知の方法を用いて、例えばサンプルを採取し、所要により該サンプル中に含まれる細胞を破壊して標的核酸を放出することにより調製され得る。この方法でPCRを使用する場合、標的配列の末端は通常公知である。LCRまたはその修正法が好ましい方法で 사용되는場合には、全標的配列は通常公知である。通常、標的配列はRNAまたはDNAのような核酸配列である。

【0119】

プライマー及びプローブの長さは変更可能であるが、プローブ配列はプライマー配列よりも低い融点を有するように選択される。よって、プライマー配列は通

常プローブ配列よりも長い。通常、プライマー配列は20～50ヌクレオチド長、より一般的には20～30ヌクレオチド長の範囲である。好ましいプライマー配列は、通常20ヌクレオチド以上の長さを有する。一般的なプローブは10～25ヌクレオチド長、より一般的には15～20ヌクレオチド長の範囲である。好ましいプローブ配列は一般的には15ヌクレオチド以上の長さを有する。

【0120】

或いは、プローブを“ネステッドPCR”として公知の方法を介して標的配列を増幅する際に関与し得る。ネステッドPCRでは、プローブは増幅に通常使用される第1及び第2プライマーと同様の特性（例えば、長さ、融点等）を有し、そのまま増幅反応においてプライマーとして作用し得る。通常、ネステッドPCRでは、第1対のプライマー（ P_1 及び P_2 ）を使用して一次延長産物を形成する。一次プライマーの1つ（例えば、 P_1 ）は場合により（第1反応性対のメンバーに結合した）捕捉プライマーであり得、他方の一次プライマー（ P_2 ）はそうではない。その後、5'末端に捕捉タイプの標識（例えば、第2反応対のメンバー）または検出標識を有し得るプローブ（ $P_{1'}$ ）またはプローブ（ $P_{2'}$ ）を用いて二次延長産物を形成する。前記プローブは、溶解しているならば P_1 及び P_2 の3'末端がハイブリダイズするであろうサイトの近くまたは該サイトに隣接する鋳型上のサイトに相補性であり、該サイトでハイブリダイズする。或いは、 P_1 プライマーとプローブ（ $P_{2'}$ ）または P_2 プライマーとプローブ（ $P_{1'}$ ）を用いて二次延長産物を形成してもよい（“ヘミネステッドPCR”とも称することもある）。よって、標識プライマー/プローブ組により、一次延長産物よりも短い二次延長産物が生成する。更に、この二次産物はそのサイズに基づいてまたはその標識末端により（当業者に公知の検出方法により）検出され得る。この方法では、プローブ及びプライマーは通常均等濃度で使用される。

【0121】

プライマー及びプローブを合成するための各種方法は当業界で公知である。また、標識をプライマーまたはプローブに結合させる方法も当業界で公知である。例えば、一般的なヌクレオチドホスホルアミダイト化学及びApplied Biosystems, Inc.（カリフォルニア州フォスターシティーに所在）

、DuPont（デラウェア州ウィルミントンに所在）またはMilligen（マサチューセッツ州ベッドフォードに所在）から入手可能な器具を用いて所望の核酸プライマーまたはプローブを合成することは日常的に行われている実験の範囲である。本発明のプライマーまたはプローブのようなオリゴヌクレオチドを標識するための多くの方法が記載されている。Enzo Biochemical（ニューヨーク州ニューヨーク）及びClontech（カリフォルニア州パロアルト）はプローブ標識法を記載し、市販している。例えば、一級アミンは3'-Amine-ON CPG（登録商標）（カリフォルニア州パロアルトに所在のClontech）を用いて3'オリゴ末端に結合され得る。同様に、一級アミンはAminomodifier II（登録商標）（Clontech）を用いて5'オリゴ末端に結合させ得る。前記アミンは慣用の活性化及び結合化学を用いて各種ハプテンに反応させ得る。更に、1992年6月25日に公開された国際特許出願公開第92/10506号パンフレット及び1994年3月1日に米国特許第5,290,925号明細書は、それぞれ5'及び3'末端でプローブを標識する方法を教示している。更に、1992年7月9日に公開された国際特許出願公開第92/11388号パンフレットは、末端でプローブを標識する方法を教示している。オリゴヌクレオチドを標識するための1つの公知の方法では、標識-ホスホルアミダイト試薬を作成し、この試薬をその合成中にオリゴヌクレオチドに標識を添加するために使用する。例えば、N.T.Thuongら、Tet.Letters, 29(46):5905-5908(1988)またはJ.S.Cohenらの米国特許出願第07/246,680号明細書(1989)を参照されたい。好ましくは、プローブを3'及び5'末端で標識する。

【0122】

捕捉標識はプライマーまたはプローブにより実施され、固相試薬特異的結合メンバーと結合対を形成する特異的結合メンバーであり得る。勿論、プライマーまたはプローブそれ自体は捕捉標識として機能し得ると解される。例えば、固相試薬特異的結合メンバーが核酸配列の場合には、プライマーまたはプローブの相補的部分に結合して前記プライマーまたはプローブを固相に固定化させるように選

択される。プローブそれ自体が結合メンバーとして機能する場合には、プローブが一本鎖アンブリコンメンバーと相補的でない配列、すなわち“テール”を含むと当業者は認識している。プライマーそれ自体が捕捉標識として機能する場合、プローブがプライマー配列に対して完全に相補的でないようにプローブが選択されているのでプライマーの少なくとも一部が固相上の核酸とハイブリダイズしない。

【0123】

通常、プローブ/一本鎖アンブリコンメンバー複合体は、異種イムノアッセイを実施するように通常使用される方法を用いて検出され得る。好ましくは、この実施態様では、検出は市販されているAbbott LCx（登録商標）（イリノイ州アボットパークに所在のAbbott Laboratories）で用いられるプロトコルに従って実施される。

【0124】

当業界で公知の他の有用な方法には、液中ハイブリダイゼーション並びにドット及びスロットブロットハイブリダイゼーションプロトコルが含まれる。場合によりサンプル中に存在する標的核酸の量は、ハイブリダイズした断片の放射能を当業界で公知の一般的方法を用いて測定することにより定量化され得る。

【0125】

III. ワクチン

ワクチンは、USタイプ及び/またはUSサブタイプ特異的タンパク質配列または前記タンパク質配列に結合する抗体に基づいて1つ以上の免疫原ポリペプチドから作成され得ると考えられる。更に、ワクチンは、死滅させた、生の弱毒化したUSタイプもしくはUSサブタイプE型肝炎ウイルス、またはUSタイプもしくはUSサブタイプE型肝炎ウイルス特異的抗原を発現する異種宿主細胞（例えば、ワクチンウイルス）を含む生の組み換えワクチンを含み得ると考えられる。

【0126】

ポリペプチドベースのワクチンの場合、ポリペプチドは少なくとも1つのエピトープを規定しなければならない。しかしながら、ワクチンは1つ以上のポリペ

ブチド鎖により規定される複数の異なるエピトープを含み得ると考えられる。更に、たとえ非構造タンパク質及び構造タンパク質が中和抗体を産生しなくても、前記タンパク質はウイルス病原性に対して保護を与え得ると考えられる。以上を考慮して、USタイプもしくはUSサブタイプウイルスに対する多価ワクチンは1つ以上の構造タンパク質及び/または1つ以上の非構造タンパク質を含み得る。前記免疫原性エピトープを組合せて、すなわち組み換えタンパク質、合成ペプチド及び/またはビリオンから単離されるポリペプチドの混合物として使用され得、同時にまたは時間をずらして一緒に投与してもよい。

【0127】

活性成分として少なくとも1つの免疫原性ペプチドを含むタンパク質またはペプチドベースのワクチンの製造方法は当業界で公知である。通常、前記ワクチンは溶液または懸濁液のような注射剤として製造される。調製物を乳化させてもよく、タンパク質をリポソームにカプセル化してもよい。活性な免疫原性成分を該活性成分と相容性の医薬的に許容され得る賦形剤と混合してもよい。好適な賦形剤には、水、食塩液、デキストロース、グリセロール、エタノールまたはその組合せが含まれるが、これらに限定されない。ワクチンは少量の助剤（例えば、湿潤・乳化剤、pH緩衝剤及び/またはワクチンの効果を高めるアジュバント）を含み得る。例えば、前記アジュバントには、2%スクアレン/ツイーン80（登録商標）エマルション中の水酸化アルミニウム、N-アセチル-ムラミル-L-スレオニル-D-イソグルタミン（Thr-DMP）、N-アセチル-ノルムラミル-L-アラニル-D-イソグルタミン（CGP 11687、nor-MDPとも称する）、N-アセチル-ムラミル-L-アラニル-D-イソグルタミニル-L-アラニン-2-(1', 2'-ジパルミトイル sn-グリセロ-3-ヒドロキシホスホリルオキシ)-エチルアミン（CGP 19835A、MIT-PEとも称する）及びRIBI（MPL+TDM+CWS）が含まれる。アジュバントの有効性は、調査しようとする各種アジュバントをも含むワクチンの形態でUSタイプもしくはUSサブタイプ特異的抗原配列を含む免疫原性ポリペプチドを投与すると生ずる前記ポリペプチドに対する抗体の量を測定することにより調べられ得る。

【0128】

前記ワクチンは通常静注または筋注により投与される。他の投与モードに適した別の製剤には座剤が含まれ、場合により経口製剤が含まれる。座剤の場合、慣用されている結合剤及び担体にはポリアルキレングリコールまたはトリグリセリドが含まれるが、これらに限定されない。前記座剤は、約0.5～約10% (w/w)、好ましくは約1～約2% (w/w)の活性成分を含有する混合物から形成され得る。経口製剤は賦形剤をも含み得、その賦形剤としてはマンニトール、ラクトース、スターチ、ステアリン酸マグネシウム、サッカリンナトリウム、セルロース、炭酸マグネシウム等が挙げられる。前記組成物は溶液、懸濁液、錠剤、ピル剤、カプセル剤、徐放性製剤または散剤の形態をとり得、約10～約95% (w/w)、好ましくは約25～約70% (w/w)の活性成分を含有する。

【0129】

ワクチン中に使用されるポリペプチド鎖は中性または塩の形態としてワクチンに処方され得る。医薬的に許容され得る塩の例には、無機酸（例えば、塩酸またはリン酸）、有機酸（例えば、酢酸、シュウ酸、酒石酸、マレイン酸）または他の当業界で公知の酸を添加して形成される酸付加塩が含まれる。遊離カルボキシル基で形成される塩は、無機塩基（例えば、水酸化ナトリウム、水酸化カリウム、水酸化アンモニウム、水酸化鉄等）、有機塩基（例えば、イソプロピルアミン、トリメチルアミン、2-エチルアミノエタノール、ヒスチジンプロカイン）または他の当業界で公知の塩基から誘導され得る。

【0130】

ワクチンは製剤に適合し得る方法で、予防及び/または治療上有効であろう量投与される。投与量は通常1回投与あたり約5～約250 µgであるが、実際の投与量は患者の健康及び体格、患者の免疫系の抗体合成能力及び所望する保護の程度に依存する。ワクチンは1回または複数回投与スケジュールで投与され得る。複数回投与とは、免疫化の初期コースでは1～10回別々に投与し、その後免疫応答を維持及び/または補強するのに必要な時間間隔で（例えば、第二回投与の間に1～4ヶ月で）、及び患者によつては所要により数ヶ月後に1回以上投与することである。更に、投与レジメは少なくとも部分的には患者の必要性により

決定され、担当医の判断に基づき得る。

【0131】

死滅させたまたは不活化させたUSタイプもしくはUSサブタイプE型肝炎ウイルスを含有するワクチンの場合、不活化は当業界で公知であり、文献に詳しく記載されている慣用方法により実施され得る。好ましい不活化方法の例には、(i)有機溶媒、(ii)洗剤、(iii)ホルマリン及び(iv)イオン放射線の1つ以上への露出が含まれる。弱毒化ワクチン中の幾つかのタンパク質は他の公知のウイルスと交差反応し得、よってUSタイプもしくはUSサブタイプE型肝炎ウイルスとHEVファミリーの他のメンバー（例えば、ビルマまたはメキシコファミリーのメンバー）の間に共通のエピトープが存在し得、従って前記病原菌により引き起こされる1つ以上の疾患に対する保護抗体を産生し得る。好ましい投与製剤及び投与方法は当業界の文献に詳しく記載されており、従って本明細書では詳細に説明しない。考慮すべき各種因子には、ペプチドベースのワクチンに関して上で検討した1つ以上の要件が含まれる。

【0132】

弱毒化生ワクチンの場合、弱毒化ウイルスは当業界で公知であり、使用されている弱毒化方法を用いて作成され得る。簡単に説明すると、弱毒化はウイルスを低温で経代することにより、ウイルスゲノムにミスセンス突然変異または欠失を導入することにより実施され得る。好ましい投与製剤及び投与方法は当業界の文献に詳しく記載されており、従って本明細書では詳細に説明しない。考慮すべき各種因子には、ペプチドベースのワクチンに関して上で検討した1つ以上の要件が含まれる。

【0133】

弱毒化生ワクチン（ベクターワクチン）の場合、このワクチンは生きているが有害でないウイルスまたはバクテリアに抗原決定基を規定するUSタイプもしくはUSサブタイプE型肝炎特異的ポリペプチド鎖をコードする遺伝子または核酸配列を導入することにより作成され得る。生じたベクター生物はその後意図する宿主に投与され得る。通常、ワクチンが有効であるためには、ベクター生物は生存可能でなければならず、本質的に非毒性であるかまたは弱毒化表現型を有して

いなければならない。好ましい宿主生物には、ワクシニアウイルス、アデノウイルス、アデノ関連ウイルス、サルモネラまたはマイコバクテリアが含まれる。ワクチンウイルス及びマイコバクテリアの生存株はそれぞれ天然痘及び結核（BCG）ワクチンの形態でヒトに安全に投与されてきた。また、前記ワクチンは外来タンパク質を発現せず、有毒表現型に殆どまたは全く変換されないことが判明している。ベクターワクチンは複数の外来遺伝子または核酸配列を有し得、よって各種の所定抗原決定基に対して同時にワクチン接種することができる。好ましい投与製剤及び投与方法は当業界の文献に詳しく記載されており、従って本明細書では詳細に説明しない。

【0134】

IV. 抗-USタイプもしくはUSサブタイプE型肝炎ウイルス活性を有する分子の同定

特異的HEV USタイプ配列の発見から、当業者はHEVゲノムのORF1タンパク質によりコードされるHEV USタイプ特異的タンパク質（例えば、ヘリカーゼ、メチルトランスフェラーゼまたはプロテアーゼタンパク質）の活性を不活化または低下させる分子を同定することができる。HCVプロテアーゼを阻害する分子を同定するためのプロトコルの例は、援用により本明細書に含まれるとする米国特許第5,597,691号明細書に記載されている。前記方法はHCVプロテアーゼ阻害剤の同定に関しているが、HCVプロテアーゼ阻害剤またはHEV USタイプ配列によりコードされる他のタンパク質を同定するために同一または類似のプロトコルを使用し得ると考えられる。

【0135】

簡単に説明すると、HEVプロテアーゼ阻害剤の同定方法は以下の通りである。通常、プロテアーゼ天然基質に似ているが開裂したときに定量可能な信号を与える基質を使用する。前記信号が比色または蛍光定量手段により検出可能であることが好ましい。しかしながら、HPLC、シリカゲルクロマトグラフィー、核磁気共鳴法等のような他の方法も有用であり得る。最適基質及びプロテアーゼ濃度を決定したら、候補プロテアーゼ阻害剤を各種濃度で反応混合物に一度に添加する。好ましいアッセイ条件は、プロテアーゼがインビボで阻害される条件、す

なわち生理学的 pH、温度、イオン強度等に類似している。好適な阻害剤は、患者に有毒な副作用を与えない濃度で強いプロテアーゼ阻害を示す。プロテアーゼ活性サイトへの結合に関して競合する阻害剤は基質濃度以上の濃度を必要とし得、プロテアーゼ活性サイトに非可逆的に結合し得る阻害剤は酵素濃度のオーダーの濃度で添加され得る。

【0136】

阻害剤は、例えば H E V プロテアーゼにより認識される開裂サイトに似ている有機化合物であり得るか、または H E V プロテアーゼに特異的に結合し且つ H E V プロテアーゼの活性を不活化または低下させ得る抗体または抗体断片のようなタンパク質であり得ると考えられる。一旦同定されれば、プロテアーゼ阻害剤は各種方法、例えば静脈内、経口、筋肉内、腹腔内、気管支、鼻腔内等で投与され得る。好ましい投与ルートは阻害剤の種類に依存する。有機化合物として製造された阻害剤は、十分に吸収されるならば経口投与される（経口投与が通常好ましい）。多くの抗体または抗体誘導体のようなタンパク質ベースの阻害剤は通常非経口ルートで投与される。

【0137】

（実施例）

本発明の実施は下記実施例から十分に理解されるであろう。これらの実施例は本明細書に例示の目的のみで提示されており、本発明を決して限定するものと解釈されるべきではない。上掲または下掲する特許明細書、特許出願明細書及び発表された科学文献を含めたすべての文献が援用により本明細書に含まれるとする。

【0138】

実施例 1：症例研究

H E V 株 U S - 1 は急性肝炎を患っている患者（U S P - 1）の血清中で同定された。この患者は 62 歳の白人男性であり、発熱、腹痛、黄疸及びそう痒感を 3 週間訴えた後ミネソタ州ロチェスターの病院に入院した。カリフォルニア州サンノゼに 10 日間旅行後自宅に戻ってから 2 週間後に兆候及び症状が発現した。

【0139】

患者は過去に、中程度の腎不全を伴う常染色体優性遺伝多嚢胞性腎疾患のために腎摘出及び症候性胆石症のために腹腔鏡下胆嚢摘出術を受けていた。患者は変形性関節炎を患っており、血圧が高かった。入院の3ヶ月前からリシノプリル治療を受けていた。診察は、白人男性が圧痛性肝肥大を伴う黄疸にかかっていることを示し、羽ばたき振せんはなかった。血清アスパルテートアミノトランスフェゼ (AST)、アラニンアミノトランスフェラーゼ (ALT) 及びビリルビンレベルは入院時に著しく高く、それぞれ入院から8日及び16日目にピークに達した (図2)。リシノプリルは入院時に中止した。A型肝炎 (IgM及びIgG抗HAV)、B型肝炎 (HbsAg、IgM及びIgG抗-HBc)、C型肝炎 (抗-HCV) 及びHCV RNAについての血清検査は陰性であった。セルロプラスミン、鉄、トランスフェリン、抗-核及び抗-平滑筋抗体、毒素及び薬物スクリーニングはすべて正常であった。患者を注意深く問診したところ、エタノールは飲用していなかった。腹部超音波、CTスキャン及び内視鏡的逆行性胆管膵臓造影も正常であった。肝生検は、自己免疫、薬物またはウイルス肝炎と一致した肝細胞の顕著な核濃縮及び風船状変性を伴う重篤な急性小葉肝炎を示した。

【0140】

患者は2ヶ月以内に臨床的に完全に回復し、入院から約5ヶ月後にAST、ALT及びビリルビンは正常となった。HEVに感染する危険因子は同定されなかった。患者は10年以上米国の国外に旅行したことはなかった。発病前6週間の間に彼が外食したのはメキシコレストランと大きなファーストフードレストランチェーンのみであった。彼は処理済飲料水しか飲用しておらず、生貝も食しておらず、家畜にも接していなかった。メキシコレストランまたはファーストフードレストランの料理人は誰も、入院日の5ヶ月以内に外国に旅行したことはなく、肝炎の兆候及び/または症状も訴えていなかった。入院期間中カリフォルニアに滞在していたとかミネソタに住んでいた非-A B C肝炎患者の他の症例は地方保険局に報告されていなかった。患者がカリフォルニアに旅行中にもその後10週間の間にも家族は誰も肝炎の兆候及び/または症状を呈しなかった。カリフォルニアの6家族及び問題の期間患者と一緒にミネソタに住んでいた彼の配偶者から得た血清はEIAにより抗-HEV陰性であった。

【0141】

実施例2：HEV US-1のユニーク単離物の同定

HEVの存在をHEVプライマー配列を用いてRT-PCRにより調べた。簡単に説明すると、核酸を患者USP-1の血清25 μ lから上記したように(Schlauder, J. Virological Methods, 46:81-89 (1995))単離した。エタノール沈降核酸を水で処理したジエチルピロカーボネート(DEPC)3 μ lに再懸濁させた。

【0142】

cDNA合成及びPCRはPerkin-Elmer(コネチカット州ノーウオークに所在)のGeneAmpRNA PCRキットを製造業者の指示に従って用いて実施した。cDNA反応物10 μ l毎にRNA(1 μ l)を鋳型として使用した。cDNA合成は4 μ Mの最終濃度まで添加した特定プライマーを用いて開始した。その後のcDNAの増幅を0.8~1.0 μ Mの最終濃度まで添加したオリゴヌクレオチドを用いて開始した。PCRを40サイクル(94 $\times$ 20秒、55 $\times$ 30秒、72 $\times$ 30秒)、その後72 $\times$ 3分の延長サイクルを実施した。次いで、最初のPCR反応物(2 μ l)をネステッドPCRプライマーを用いる第2ラウンドの増幅のための鋳型として使用した。PCRはPerkin-ElmerのGeneAmp PCRキットを製造業者の指示に従って用いて実施した。簡単に説明すると、プライマーを1 μ Mの最終濃度まで添加した。第1組の実験では3組のプライマーを用いた。2組はビルマ株及びメキシコ株由来の配列に基づくORF1の5'末端からであった。1組はメキシコ株配列に基づくORF1の3'末端からであった。使用した3組のプライマーは以下の通りであった。

【0143】

【表1】

プライマー組 1

プライマー	配列	配列番号
5'-ORF1-メキシコプライマー C375M	CTGAACATCCCGGCCGAC	1
PCR プライマー A1-350M	AGAAAGCAGCGATGGAGGA	2
PCR プライマー S1-34M	GCCCACCAGTTCATTAAGGCT	3
ネステッド PCR プライマー A2-320M	TCATTAATGGAGCGTGGGTG	4
ネステッド PCR プライマー S2-55M	CCTGGCATCACTACTGCTAT	5

プライマー組 2

プライマー	配列	配列番号
5'-ORF1-ヒルマ cDNA プライマー C375	CTGAACATCACGCCCAAC	6
PCR プライマー A1-350	AGGAAGCAGCGGTGGACCA	7
PCR プライマー S1-34	GCCCATCAGTTTATTAGGGC	8
ネステッド PCR プライマー A2-320	TCATTTATTGAGCGGGGATG	9
ネステッド PCR プライマー S2-55	CCTGGCATCACTACTGCTAT	10

プライマー組 3

プライマー	配列	配列番号
3'-ORF1-メキシコ cDNA プライマー M1PR6	CCATGTTCCACACCGTATTCCAGAG	11
PCR プライマー S4294M	GTGTTCTACGGGGATGCTTATGACG	12
ネステッド PCR プライマー M1PF6	GACTCAGTATTCTCTGCTGCCGTGG	13
ネステッド PCR プライマー A4556	GGCTCACCAGAATGCTTCTTCCAGA	14

【0144】

5'-ORF1ヒルマプライマーはSchlauderら, Lancet, 341:678(1993)に記載されている。プライマーM1PR6及びM1PF6はMcCautlandら, J. Virological Methods, 35:331-342(1991)に記載されている。PCR産物をアガロースゲル電気泳動により分離し、臭化エチジウム染色後UV照射することにより可視化した。サザンブロットをニトロセルロースフィルターに移した後生じたPCR産物を放射性同位元素標識したプローブにハイブリダイズした。

【0145】

放射性同位元素標識したプローブは、Qiagen(カリフォルニア州チャッ

ツワースに所在)のQ I A E X ゲル抽出精製により精製したP C R 産物から作成した。放射性同位元素標識はS t r a t g e n e (登録商標)(カリフォルニア州ラ・ホーヤに所在)P r i m e - I t I I キットを製造業者の指示に従って用いて取り込んだ。フィルターをA m e r s h a m (イリノイ州アーリントンハイツに所在)のR a p i d - h y b 緩衝液において3 ~ 5 時間プレハイブリダイズした後、42 °C でF a s t - P a i r ハイブリダイゼーション溶液において10 ~ 200 c p m / c m²で15 ~ 25 時間ハイブリダイズした。次いで、フィルターをS c h l a u d e r 著, J . V i o l . M e t h o d s , 37 : 189 - 200 (1992)に記載されているように洗浄した。プローブ化フィルターのリンイメージをM o l e c u l a r D y n a m i c s P h o s p h o r i m a g e r 425 E (カリフォルニア州サニーベールに所在)を用いて得た。

【0146】

臭化エチジウム染色バンドをO R F 1 の5' 末端からのプライマーを用いて検出した。しかしながら、メキシコ株に基づくプライマーのみにより266 塩基対の予想サイズを有するネステッド産物が生じた。ビルマ様株(同一性>90%)感染患者由来のプローブへのハイブリダイゼーションにより、ビルマポジティブコントロールからの信号に比して患者U S P - 1 由来産物に対して非常に弱いハイブリダイゼーション信号が生じた。これらの結果は、この単離物がビルマ単離物に密接に関連していなかったことを初めて示すものであった。メキシコ株からプローブは入手できなかった。

【0147】

上記結果を確認するために、患者U S P - 1 の別の血清アリコートからR N A を抽出した。R T - P C R を上記したように5' - O R F 1 - メキシコプライマー(配列番号1 ~ 5)を用いて実施した。アガロース電気泳動及び臭化エチジウム染色後、各サンプルにおいて342 b p 産物が可視化された。P C R 産物をQ i a g e n (カリフォルニア州チャッツワースに所在)のQ I A E X I I アガロースゲル抽出キットを用いてアガロースゲルから抽出し、N o v a g e n (ウィスコンシン州マディソンに所在)のp T 7 B l u e T - ベクタープラスミドにクローン化した。クローン化産物をS E Q U E N A S E V E R S I O N 2

．0配列決定キット（オハイオ州クリーヴランドに所在のU S B）を製造業者の指示に従って用いて配列決定した。

【0148】

後者の2つのサンプルの産物から得たヌクレオチド配列は同一であり、配列番号15に示す。これらの結果から、ビルマ株及びメキシコ株由来のcDNAプライマー及びプライマーS1のみにより患者USP-1サンプルから臭化エチジウムで染色可能な産物が生じたことが分かる。メキシコ株に基づくネステッドプライマーS2及びA2だけが所望サイズを有する臭化エチジウムで染色可能な産物を生成した。

【0149】

HEV US-1単離物とHEVの他の公知の単離物の関連度を調べるために、ウィスコンシン州マジソンに所在のScience Drive 575に所在のGenetics Computer Group, Inc. から入手し得るWisconsin Sequence Analysis Package（バージョン9）のプログラムGAPを用いてヌクレオチド配列及びアミノ酸配列のアラインメントを実施した。前記プログラムは、アラインさせた2つの配列間の%として表す類似度及び同一度を計算するためにNeedleman及びWunschのアルゴリズム（J. Mol. Biol., 48:443-453（1970））を使用している。ギャップ形成及びギャップ延長ペナルティーは核酸配列アラインメントについてそれぞれ50及び3.0であり、アミノ酸配列比較に対してそれぞれ12及び4であった。

【0150】

ビルマ株及びメキシコ株由来の2つの“原型”HEV単離物の完全ヌクレオチド配列及びアミノ酸配列、並びに分析のために使用した他の配列はGenBankから入手し、それぞれの受託番号を下表1に示す。これらの配列の各々は援用により本明細書に含まれるとする。

【0151】

【表2】

表1

単離物	Genbank 受託番号
メキシコ (M1)	M74506
ビルマ (B1)	M73218
パキスタン (P1)	M80581
中国 (C4)	D11093

(配列番号89の残基1～303に相同の)HEV US-1の303塩基対配列をメキシコ株、ビルマ株、パキスタン株及び中国株で同定した相同領域と比較した。同一%を下表2に要約する。

【0152】

【表3】

表2：5'末端ORF1産物由来の303核酸の範囲の同一性

	US-1	メキシコ	ビルマ	パキスタン
メキシコ	77.2			
ビルマ	74.9	83.2		
パキスタン	75.9	83.2	95.7	
中国	75.9	83.5	95.7	97.4

【0153】

表2の結果から、USP-1単離物由来のORF1の5'末端からの断片がHEVの他の公知の単離物に対して約74.9～約77.2%核酸同一であることが分かる。これは、原型メキシコ単離物とビルマ単離物間の同一性(83.2%)よりも低かった。これらの結果は、産物が多分今までに同定されていなかったHEVのユニークな単離物に由来していたことを示している。

【0154】

実施例3：HEV US-1のゲノム延長及び配列決定

上記実施例2に記載のように入手し、配列決定したクローン(配列番号15)はユニークなHEVゲノム、HEV US-1由来であった。HEV US-1ゲノムの別の領域から配列を得るために、幾つかの逆転写酵素-ポリメラーゼ連鎖反応(RT-PCR)ウォーキング実験を実施した。

【0155】

全核酸を、実施例2に記載の手順（配列番号19に対してのみ）または以下の手順の1つにより抽出した。患者USP-1血清のアリコート（25 µl）を担体としての酵母tRNA（10 mg）の存在下でTotal Nucleic Acid Extraction方法を製造業者（United State Biochemical）に従って用いて抽出した。核酸を沈降させ、RNase/DNase非含有水（3.75 µl）中に再懸濁させた。或いは、全RNAを血清（100 µl）からTOTALLY RNA単離キットを製造業者（Ambion, Inc.）が推奨するように用いて単離した。生じたRNAをDNaseで処理し、S.N.A.P. Total RNA単離キット（カリフォルニア州サンジェゴに所在のInvitrogen）の試薬を用いてカラム精製した。その後、RNAを3M 酢酸ナトリウム（0.1 容量）、担体としてのペレットポイント（Novagen）（20 µl）及びエタノール（2 容量）を用いて沈降させた。RNAペレットをDEPC処理水（50 µl）に溶解させた。

【0156】

RT-PCRをGeneAmp RNA PCRキットを製造業者の指示（Perkin-Elmer）に従って用いて実施した。配列番号19の単離物を除いて、総容量25 µlでcDNA合成を開始するためにランダムヘキサマーを用いた。配列番号19の単離物に対しては、上記実施例2に記載したようにプライマーPA2-5560（配列番号16）を用いて特異的に開始させたcDNAを利用した。US1-ギャップを、12.5 µlの血清等価物から抽出したRNA、プライマーUS1ギャップ-a0.5（配列番号46）及びSuperscript II（3' RACKキット：GIBCO BRL）を用いて作成した特異的プライマーcDNAで作成した。PCRを全反応容量の1/5（10 µl 反応物では2 µl、25 µl 反応物では5 µl等）を含むcDNAを用いて実施した。標準PCRを2 mMのMgCl₂及び0.5 ~ 1.0 Mの各プライマーの存在下で実施した。配列番号33及び41の単離のために製造業者の指示に従って修飾反応物に1 × PCR緩衝液及び20% Q Solution（Qiagen）を含めた。反応には2つのHEVコンセンサスプライマー（表3）；1つのHEVコンセンサスプライマー及び1つのHEV US-1特異的プライマー

(表4) ; 2つのHEV US - 1特異的プライマー(表5) ; 1つのHEV US - 1特異的プライマー及び1つのHEV US - 2(実施例5参照)特異的プライマー ; または2つのHEV US - 2特異的プライマー(表7)を使用した。反応物を次のように熱サイクルにかけた。

【0157】

配列番号19、24、27、30、33、41、44、60、64、68、73、78及び83をタッチダウンPCRにより得た。増幅には、94 × 30秒、55 × 30秒(-0.3 / サイクル) 及び72 × 1分間の43サイクルを含めた。この後に、94 × 30秒、40 × 30秒及び72 × 1分の10サイクルを実施した。配列番号38、49、52及び55の場合、94 × 30秒、55 × 30秒及び72 × 1分間の35サイクルを含めた。すべての増幅の前に94 × 1~2分、その後に72 × 5~10分間を含めた。反応物をアガロースゲル分析前に4 で保持した。

【0158】

配列番号19の単離には、所望産物を単離するために2ラウンドのタッチダウン増幅が必要であった。ここで、第1ラウンド1 µlを第2ラウンドの25 µl反応物に添加した。第2ラウンドの増幅は反応1.1.1及び1.1.2により表3に示すヘミネステッドプライマーを使用した。配列番号24の単離には第2ラウンドの上記したネステッドタッチダウン増幅を必要とし、反応2.1.1及び2.1.2として表4に示した。配列番号38及び49の単離には、上記したように第1ラウンド1 µlを用いる25 µl反応物に用いるネステッドPCR(表5)の第2ラウンドを必要とした。配列番号60、64、68及び73の単離には、第1ラウンド1 µlを第2ラウンド反応25 µlにおいて増幅させるネステッドPCRを必要とした(表6)。配列番号78及び83の産物が2ラウンド増幅により生成した(表7)。

【0159】

PCR反応物の一部または全部に対してアガロースゲル電気泳動を0.2 mg / mlの臭化エチジウムの存在下0.8~2%アガロースTAEゲル中で実施した。産物をUV照射により可視化し、所望分子量の産物を切り取り、GeneC

leanを製造業者の指示(BIO 101, Inc.)に従って精製し、pT7-Blue T-Vectorプラスミド(Novagen) IIまたはpGEM-T Easy Vector(Promega)に製造業者の指示に従ってクローン化した。クローン化産物を実施例2に記載したように、またはABI Model 373 DNA Sequencer上で製造業者が特定するようにABI Sequencing Ready Reaction Kitを用いて配列決定した。上記実験の結果を下表3～7に示す。

【0160】

【表4】

表3

反応	プライマー1	プライマー2	大体の産物サイズ／配列番号
1.1.1	配列番号17	配列番号16	なし
1.1.2	配列番号18	配列番号16	251bp／配列番号19
1.2	配列番号19	配列番号29	168bp／配列番号30

【0161】

【表5】

表4

反応	プライマー1	プライマー2	大体の産物サイズ／配列番号
2.1.1	配列番号20	配列番号22	なし
2.1.2	配列番号21	配列番号23	899bp／配列番号24
2.2	配列番号25	配列番号26	846bp／配列番号27
2.3	配列番号31	配列番号32	424bp／配列番号33
2.4	配列番号39	配列番号40	460bp／配列番号41
2.5	配列番号42	配列番号43	235bp／配列番号44

【0162】

【表6】

表5

反応	プライマー組 P C R 1	プライマー組 P C R 2	大体の産物サイズ/ 配列番号
3.1	配列番号 34/ 配列番号 35	配列番号 36/ 配列番号 37	1186bp/配列番号 38
3.2	配列番号 45/ 配列番号 46	配列番号 47/ 配列番号 48	545bp/配列番号 49
3.3	配列番号 50/ 配列番号 51		344bp/配列番号 52
3.4	配列番号 53/ 配列番号 54		194bp/配列番号 55

【0163】

【表7】

表6

反応	プライマー組 P C R 1	プライマー組 P C R 2	大体の産物サイズ/ 配列番号
4.1	配列番号 56/ 配列番号 57	配列番号 58/ 配列番号 59	464bp/配列番号 60
4.2	配列番号 61/ 配列番号 62	配列番号 63/ 配列番号 62	433bp/配列番号 64
4.3	配列番号 65/ 配列番号 66	配列番号 65/ 配列番号 67	382bp/配列番号 68
4.4	配列番号 69/ 配列番号 70	配列番号 71/ 配列番号 72	451bp/配列番号 73

【0164】

【表8】

表7

反応	プライマー組 P C R 1	プライマー組 P C R 2	大体の産物サイズ/ 配列番号
5.1	配列番号 74/ 配列番号 75	配列番号 76/ 配列番号 77	334bp/配列番号 78
5.2	配列番号 79/ 配列番号 80	配列番号 81/ 配列番号 82	413bp/配列番号 83

【0165】

ゲノムの3'末端で配列を得るために、増幅はGIBCO BRLの3' RACEシステムを製造業者の指示に従って使用した。HEV株のように、HRV-US-1ゲノムの3'末端はメキシコ株、ビルマ株及びパキスタン株と同様にポリアデノシン尾を含むと想定した。血清50 μ lの等価物から上記のように抽出したRNAを製造業者から供給された配列番号84のオリゴdTアダプタープライマー5'-GGCCACGCGTCGACTAGTACTTTTTTTTTTTTTTTTTT-3'を用いて逆転写した。第1ラウンドのPCRでは、AUA Pプライマー供給5'-GGCCACGCGTCGACTAGTAC-3' (配列番号85) 及び0.2 mM最終濃度のHEV-US特異的プライマー (表8) を推奨される濃度のPCR緩衝液、MgCl₂及びcDNAと共に用いた。増幅には94 $\times$ 30秒、55 $\times$ 30秒及び72 $\times$ 1時間の35サイクルを含めた。増幅の前に94 $\times$ 1分のインキュベーション、増幅の後に72 $\times$ 10分の延長を行った。増幅の第2ラウンドでは、50 μ lの反応に1 μ lの第1ラウンドを使用した。PCR緩衝液は2 mMのMgCl₂及び0.5 mMの各プライマーで1 $\times$ 最終濃度であった。プライマーをAUA Pプライマー及びHEV-US-1特異的プライマー (表8) を用いてヘミネステッドした。増幅条件は第1ラウンドと同一であった。産物を上記のようにアガロースゲル電気泳動により分析、クローン化し、配列決定した。

【0166】

【表9】

表8

反応	プライマー組 PCR 1	プライマー組 PCR 2	大体の産物サイズ／配列番号
8.1	配列番号 86／配列番号 85	配列番号 87／配列番号 85	960bp／配列番号 88

【0167】

上記産物から得た配列を上表3～8に記載し、ゲノム (配列番号15) の5'末端近くの最初のPCR産物をGCGパッケージ (ウィスコンシン州マジソンに所在のGenetics Computer Group, バージョン9) のブ

ログラムを用いてコンティグにアセンブリし、コンセンサス配列を決定した。アセンブリしたコンティグの概略を図3に示す。HEV US-1ゲノムは7202bp長であり、そのすべての配列が決定された(配列番号89)。この配列を3つのオープンリーディングフレームに翻訳した。前記オープンリーディングフレームの2つを配列番号90に示す(第3のORFはヌクレオチド位置5094~5462にあるが、他の2つのORFと重複するので配列番号90に示すことはできない)。生じた翻訳物(ORF1、ORF2及びORF3)をそれぞれ配列番号91、配列番号92及び配列番号93に示す。

【0168】

実施例4：HEV US-2のユニーク単離物の同定

HEV EIA試験でIGGクラス抗体について検査した急性肝炎を患っている米国患者も、US-1株特異的ELISAを用いて陽性と判定された。急性肝炎と診断されたこの患者(US-2)は黄疸及び疲労を訴えて入院した62歳男性であった。最初の臨床検査で、ALTは1270U/Lであった(正常値0~40U/L)。当地域で最近A型肝炎ウイルスが大発生しているので、この患者もHAVに感染していると疑われた。しかしながら、抗-HAV IgM検査のHAVAB-M EIA(Abbott Laboratories)は、B型肝炎ウイルス及びC型肝炎ウイルスに対する血清マーカーに関する検査と同様に陰性であった。この患者の病歴によると、発病の数週間前にメキシコのカンクンを訪問していた。

【0169】

その後、患者からのサンプルをHEV US-1特異的PCRプライマーを用いるPCR増幅によりHEV特異的配列の存在について分析した。RNAを実施例2に記載のUltraspexを用いて抽出した。ランダムプライマーcDNA合成を実施例3に記載のように実施し、PCRをHEV US-1特異的プライマー(配列番号94及び配列番号96)を用いて実施例2に記載の標準条件で実施した。ネステッドPCRはプライマー(配列番号95及び配列番号97)を用いて実施した。PCR産物は実施例3に記載のように配列決定した。生じたPCR産物の配列を配列番号98に示す。実施例2のようにGAP分析して、ヌク

レオチド配列（配列番号98）はHEV US - 1からの対応または相同領域と95%同一であった。

【0170】

実施例5：HEV US - 2のゲノム延長及び配列決定

実施例4で得、配列決定したクローン（配列番号98）はHEV US - 1に最も密接に関連しているHEV単離物から誘導した。HEV US - 2ゲノムの別の領域を得るために、幾つかのRT - PCRウォーキング実験を実施例3に記載のように実施した。

【0171】

RNAをTotal Nucleic Acid Extraction方法（United States Biochemical）を用いて抽出した。逆転写物をGeneAmp RNA PCRキット（Perkin - Elmer）を用いてランダムプライムした。標準PCRを2mMのMgCl₂及び0.5 ~ 1.0 μMの各プライマーの存在下で実施した。配列番号129、141及び146の単離のために修飾反応物には1×PCR緩衝液及び20% Q Solution（Qiagen）を含めた。反応には2つのHEV US - 1特異的プライマー（表9）；1つのHEV US - 1特異的プライマー及び1つのHEVコンセンサスプライマー（表10）；1つのHEV US - 2特異的プライマー及び1つのHEVコンセンサスプライマー（表11）；2つのHEV US - 2特異的プライマー（表12）；2つのビルマ、メキシコ及びUS由来コンセンサスプライマー（以下に記載する、表13）を使用した。

【0172】

配列番号101、102、105、108、110、113、117、120、124、149及び151に示す産物はタッチダウンPCRにより得た。増幅には、94 × 30秒、55 × 30秒（- 0.3 / サイクル）及び72 × 1分の43サイクルを含めた。増幅の後に、94 × 30秒、40 × 30秒及び72 × 1分の10サイクルを含めた。配列番号129、132、136、141及び146を増幅するためには94 × 30秒、55 × 30秒及び72 × 1分の35サイクルを含めた。すべての増幅において、その前に94 × 1 ~

2分、その後72 × 5 ~ 10分を含めた。この反応物を4 で保持してから、アガロースゲル分析にかけた。大抵の産物の単離には、表9 ~ 13に示すネステッドまたはヘミネステッドPCRの第2ラウンドを必要とした。これらの反応では、PCR1産物1 µlをPCR2反応混合物25 ~ 50 µlに添加し、生じた混合物をPCR1と同様のサイクルにかけた。

【0173】

上記実施例3に記載のように反応物を分析し、産物をクローン化し、配列決定した。これらの実験の結果を表9 ~ 13に示す。

【0174】

【表10】

表9

反応	プライマー組 PCR1	プライマー組 PCR2	大体の産物サイズ/ 配列番号
7.1	配列番号99/ 配列番号100		331bp/配列番号101
7.2	配列番号34/ 配列番号35	配列番号36/ 配列番号37	1186bp/配列番号102
7.3	配列番号103/ 配列番号104		130bp/配列番号105
7.4	配列番号106/ 配列番号107	配列番号39/ 配列番号107	564bp/配列番号108
7.5	配列番号86/ 配列番号109	配列番号87/ 配列番号109	678bp/配列番号110

【0175】

【表11】

表10

反応	プライマー組 PCR1	プライマー組 PCR2	大体の産物サイズ/ 配列番号
8.1	配列番号111/ 配列番号112		580bp/配列番号113
8.2	配列番号114/ 配列番号116	配列番号116/ 配列番号115	734bp/配列番号117

【0176】

【表12】

表11

反応	プライマー組 PCR1	プライマー組 PCR2	大体の産物サイズ/ 配列番号
9.1	配列番号118/ 配列番号119		483bp/配列番号120
9.2	配列番号121/ 配列番号122	配列番号121/ 配列番号123	431bp/配列番号124
9.3	配列番号125/ 配列番号126	配列番号127/ 配列番号128	1020bp/配列番号129

【0177】

【表13】

表12

反応	プライマー組P CR1	プライマー組 PCR2	大体の産物サイズ/ 配列番号
10.1	配列番号130/ 配列番号131		407bp/配列番号132
10.2	配列番号133/ 配列番号134	配列番号135/ 配列番号134	547bp/配列番号136
10.3	配列番号137/ 配列番号138	配列番号139/ 配列番号140	903bp/配列番号141
10.4	配列番号142/ 配列番号143	配列番号144/ 配列番号145	503bp/配列番号146

【0178】

【表14】

表13

反応	プライマー組	大体の産物サイズ/ 配列番号
11.1	配列番号147/配列番号148	418bp/配列番号149
11.2	配列番号150/配列番号126	197bp/配列番号151

【0179】

ゲノムの3'末端で配列を得るために、増幅は実施例3に記載されているよう

にGIBCO BRLの3' RACEシステムを製造業者の指示に従って用いた。cDNAを配列番号84を用いて作成した。PCR1はプライマー（配列番号150及び配列番号85）を用いた。PCR2プライマーは配列番号152及び配列番号85であった（反応12.1）。生じた産物は901bp（配列番号153）であった。

【0180】

HEV US-2ウイルスゲノムの5'末端にある新規配列を逆PCR（M, Zeiner及びU. Gehring, Biotechniques, 17:1051-1053（1994））により単離した。USP-1及びUSP-2からの血清は十分に入手できないので、（下記実施例に記載の）HEV US-2感染マカクの糞材料をソース材料として選択した。462ヌクレオチドの産物を、プライマー（配列番号154、155、156及び157）を用いて実施例3に記載のように抽出、逆転写及びPCR増幅したRNAを用いて超可変/富プロリンヒンジ領域内でマカク糞材料から増幅させた。この産物（配列番号158）はHEV US-2配列と100%同一であった。従って、マカク糞からのHEVゲノムの5'末端で同定された配列はHEV US-2ゲノムの5'末端を正確に表さなければならないと考えられる。全核酸を上記した10%糞懸濁液200µlから抽出した。HEV US特異的プライマー（配列番号159）を用いる逆転写反応を、（上掲のM. Zeiner及びU. Gehring, Biotechniquesに記載されている）BMBから入手したキットを用いて実施した。但し、核酸を70℃で5分間変性し、氷上に保存した後RT反応を開始した。二本鎖環状cDNAの作成は上掲のM. Zeiner及びU. Gehring, Biotechniquesに記載されているように実施した。生じた環状cDNA分子をその後のPCR反応の鋳型として使用した。第1のPCR反応（PCR1）で用いたプライマーを配列番号160及び161に示す。第2PCR反応（PCR2）で用いたネステッドプライマーを配列番号162及び163に示す。

【0181】

PCR2（反応13.1）からの産物をpGEM-EasyTベクター（Pr

omega) にクローン化し、Applied Biosystems 3730 自動シーケンサーを用いて配列決定した。221ヌクレチドの1つの産物は適当なプライマー及び公知のHEV US-2配列の上流63ヌクレオチドを同定するHEV US-2配列を有するものとして同定された。別のクローンを適当なプライマー及び前記の新規配列の一部を用いて同定した。100 µlのUSP-2血清または100 µlの10%糞懸濁液由来のRNAに対して配列番号163及び164に示す配列をプライマーとして用いて実施したプライマー延長実験では、この配列の長さをうまく確認できなかった。ビルマ様単離物の5' NTR配列に対する63ヌクレオチドの1対比較で94%以上の同一性を示し、これがHEV US-2の真の配列であることが示唆された。

【0182】

本実施例に記載の産物から得た配列及び実施例4に記載の配列をGCGパッケージ(ウィスコンシン州マジソンに所在のGenetics Computer Group, バージョン9)のプログラムを用いてコンティグに集合し、コンセンサス配列を決定した。集合したコンティグを図4に概略的に示す。HEV US-2株のゲノムは7277 bp長であり、そのすべての配列が決定されており、配列番号164に示す。この配列を配列番号165に示す3つのオープンリーディングフレームに翻訳し、ORF1及びORF2配列の翻訳産物のみを示す(第3のORFはヌクレオチド位置5159~5527に存在するが、他の2つのORFと重複しているので配列番号165に示すことができない)。ORF1、ORF2及びORF3配列の生じた翻訳物をそれぞれ配列番号166、167及び168に示す。

【0183】

実施例6：配列比較

ウイルスの関連度に関する情報は通常、ヌクレオチド及び推定アミノ酸配列のアラインメントのような比較を実施することにより得ることができる。HEVのUS単離物(例えば、HEV US-1及びHEV US-2)の配列とHEVの他の単離物の対応配列のアラインメントにより、配列間の類似度及び同一度が定量的に評価される。通常、2つのアミノ酸配列の類似性は、アラインメント中

のアミノ酸対の側鎖間の類似度に基づいて算出される。類似度はアミノ酸側鎖の物理 - 化学的特性、すなわちサイズ、形、電荷、水素結合能及び化学反応性に基づく。従って、類似のアミノ酸は類似の物理 - 化学的特性を有する側鎖を持っている。2つのアラインさせたアミノ酸またはヌクレオチド配列間の同一性の計算は、通常アラインメント中のアミノ酸またはヌクレオチドの同一対の数をカウントし、この数をアラインメント中の配列の長さで割る算術計算である。2つのアラインさせたヌクレオチド配列間の類似性の計算には、時にはアラインメント中の異なる位置の対比（マッチド）ヌクレオチド間のトランジション及びトランスバージョンについて異なる値を使用する。しかしながら、ヌクレオチド配列の対間の類似スコア及び同スコアの大きさは通常非常に似通っており、すなわち1～2%の範囲内である。

【0184】

類似度及び同一度は、Wisconsin Sequence Analysis PackageのプログラムGAP（バージョン9）を用いて決定した。ギャップ形成及びギャップ延長のペナルティーは核酸配列の場合それぞれ50及び3.0であり、アミノ酸配列比較の場合それぞれ12及び4であった。

【0185】

上記したように、最初の5'末端ORF1クローンとHEVの他の単離物は部分的に同一であり、これはHEVのユニーク単離物のためにHEV感染が患者USP-1に関連するという説を裏付けるものである。この単離物とHEVの他の公知の単離物の関連度をより大規模に調べるために、延長したヌクレオチドと推定アミノ酸間のアラインメントを実施した。

【0186】

US単離物の相互間及びUS単離物とHEVの公知の変異体の間の関連性を調べるために、HEV US-1、HEV US-2及び（公開されているデータベースから得た、表14参照）10個の他の完全長HEVゲノムのヌクレオチド及びアミノ酸の1対比較を上記したように実施した。

【0187】

【表15】

表14

単離物	Genbank 受託番号
メキシコ (M1)	M74560
ビルマ (B1)	M73218
ビルマ (B2)	D10330
パキスタン (P1)	M80581
中国 (C1)	D11092
中国 (C2)	L25547
中国 (C3)	M94177
中国 (C4)	D11093
インド (I1)	X98292
インド (I2)	X99441

US - 1、US - 2、B1、B2、I2、C1、C2、C3、P1、C4及びI1株の全ゲノムの範囲のヌクレオチド同一性を表15に示す。ORF1、ORF2及びORF3のヌクレオチド同一性をそれぞれ表16～18に示す。表17及び18には、GenBank受託番号AF011921として入手可能な最近単離されたブタ (S1) 配列との比較も示す。

【0188】

【表16】

表15：ゲノムの範囲のヌクレオチド同一性

	US-1	US-2	B1	B2	I2	C1	C2	C3	P1	C4	I1
US-2	92.0										
B1	73.9	74.0									
B2	73.8	74.0	98.5								
I2	73.5	73.8	96.1	95.4							
C1	74.2	74.3	93.9	93.4	92.3						
C2	74.2	74.3	93.5	93.0	92.0	98.7					
C3	74.1	74.3	93.7	93.0	92.0	98.2	98.7				
P1	74.1	74.1	93.6	92.8	92.0	98.2	98.8	98.3			
C4	73.7	73.9	94.5	94.1	92.7	97.1	97.2	96.8	96.7		
I1	74.4	74.4	93.5	93.0	92.2	93.8	94.0	93.8	93.9	93.5	
M1	73.7	74.5	75.9	75.7	75.0	75.9	75.9	75.9	76.1	75.7	75.7

【0189】

【表17】

表 16 : ORF 1 の範囲のヌクレオチド同一性

	US-1	US-2	B1	B2	I2	C1	C2	C3	P1	C4	I1
US-1											
US-2	92.0										
B1	71.7	71.6									
B2	71.7	71.8	98.6								
I2	71.2	71.5	95.7	95.1							
C1	72.1	72.1	93.5	93.1	91.8						
C2	72.2	72.3	93.1	92.7	91.5	98.6					
C3	71.9	72.2	93.3	92.8	91.4	98.1	98.7				
P1	72.2	72.1	93.1	92.6	91.4	98.2	99.0	98.4			
C4	71.5	71.7	94.6	94.4	92.3	96.7	98.8	96.3	96.4		
I1	72.3	72.3	93.2	92.8	91.5	93.6	94.0	93.7	93.9	93.3	
M1	72.0	72.6	73.6	73.5	72.5	73.7	73.8	73.8	73.9	73.4	73.5

【0190】

【表18】

表 17 : ORF 2 の範囲のヌクレオチド同一性

	US-1	US-2	B1	B2	I2	C1	C2	C3	P1	C4	I1	M1
US-1												
US-2	92.2											
B1	79.2	79.6										
B2	86.4	79.4	98.5									
I2	79.0	79.5	99.2	98.4								
C1	79.3	79.5	94.4	98.4	98.4							
C2	79.2	79.4	94.3	97.8	97.8	98.9						
C3	79.3	79.4	94.4	97.8	97.8	98.9	98.4					
P1	79.0	79.3	93.8	98.1	98.7	99.7	99.2	99.2				
C4	78.8	79.3	94.0	97.8	97.8	98.9	98.4	98.4	97.4			
I1	79.4	79.7	94.1	97.6	97.3	97.9	97.0	94.0	93.7	93.9		
M1	78.0	79.3	81.1	90.1	98.5	90.6	90.1	81.0	81.4	90.3	90.3	
S1	92.0	98.9	79.8	84.6	85.4	85.4	85.1	80.2	80.1	84.8	85.1	84.6

【0191】

【表19】

表 18 : ORF 3 の範囲のヌクレオチド同一性

	US-1	US-2	B1	B2	I2	C1	C2	C3	P1	C4	I1	M1
US-1												
US-2	96.2											
B1	87.0	86.6										
B2	86.4	86.3	99.2									
I2	86.4	86.9	97.8	99.2								
C1	87.3	86.3	99.2	98.4	98.4							
C2	86.4	86.1	98.1	97.3	97.8	98.9						
C3	86.7	85.6	98.1	97.3	97.8	98.9	98.4					
P1	87.0	86.6	98.9	98.1	98.7	99.7	99.2	99.2				
C4	86.2	85.8	98.1	97.6	97.8	98.9	98.4	98.4	99.2			
I1	86.4	86.6	97.8	97.6	97.6	97.9	97.0	97.0	97.8	97.8		
M1	84.6	85.2	87.8	90.1	89.5	90.6	90.1	90.1	90.9	90.3	90.3	
S1	94.9	96.7	85.1	84.6	85.4	85.4	85.1	84.8	85.6	84.8	85.1	84.6

【0192】

更に、メチルトランスフェラーゼタンパク質をコードするORF 1ヌクレオチド配列をUS - 1、US - 2、M1及びP1単離物の各々の間で比較した。HEV US - 1ゲノムのメチルトランスフェラーゼコーディング領域は配列番号89の残基1～693で示されるのに対して、HEV US - 2ゲノムのメチルトランスフェラーゼコーディング領域は配列番号164の残基36～755で示される。比較結果を表19に示す。

【0193】

【表20】

表 19 : メチルトランスフェラーゼ領域

	同一性 (%)			
	US - 1	US - 2	M1	P1
US - 1	—	93.4	77.0	75.2
US - 2	—	—	78.5	76.0
M1	—	—	—	78.8

【0194】

Yドメインタンパク質をコードするORF 1ヌクレオチド配列をUS - 1、US - 2、M1及びP1単離物の各々の間で比較した。HEV US - 1ゲノムのYドメインタンパク質コーディング領域は配列番号89の残基619～1272で示されるのに対して、HEV US - 2ゲノムのYドメインタンパク質コーデ

ィング領域は配列番号164の残基680～1334で示される。比較結果を表20に示す。

【0195】

【表21】

表20：Yドメイン

	同一性 (%)			
	US-1	US-2	M1	P1
US-1	—	94.0	79.0	77.2
US-2	—	—	79.7	76.8
M1	—	—	—	78.3

【0196】

プロテアーゼタンパク質をコードするORF1ヌクレオチド配列をUS-1、US-2、M1及びP1単離物の各々の間で比較した。HEV US-1ゲノムのプロテアーゼタンパク質コーディング領域は配列番号89の残基1270～2091で示されるのに対して、HEV US-2ゲノムのプロテアーゼタンパク質コーディング領域は配列番号164の残基1332～2153で示される。比較結果を表21に示す。

【0197】

【表22】

表21：プロテアーゼ領域

	同一性 (%)			
	US-1	US-2	M1	P1
US-1	—	91.8	65.1	64.0
US-2	—	—	65.1	63.0
M1	—	—	—	68.1

【0198】

超可変領域をコードするORF1ヌクレオチド配列をUS-1、US-2、M1及びP1単離物の各々の間で比較した。HEV US-1ゲノムの超可変領域コーディング領域は配列番号89の残基2092～2364で示されるのに対し

て、HEV US-2ゲノムの超可変領域コーディング領域は配列番号164の残基2194～2429で示される。比較結果を表22に示す。

【0199】

【表23】

表22：超可変領域

	同一性 (%)			
	US-1	US-2	M1	P1
US-1	—	83.9	40.3	50.2
US-2	—	—	45.8	49.8
M1	—	—	—	40.4

【0200】

Xドメインタンパク質をコードするORF1ヌクレオチド配列をUS-1、US-2、M1及びP1単離物の各々の間で比較した。HEV US-1ゲノムのXドメインタンパク質コーディング領域は配列番号89の残基2365～2841で示されるのに対して、HEV US-2ゲノムのXドメインタンパク質コーディング領域は配列番号164の残基2430～2906で示される。比較結果を表23に示す。

【0201】

【表24】

表23：Xドメイン

	同一性 (%)			
	US-1	US-2	M1	P1
US-1	—	91.6	72.5	71.3
US-2	—	—	72.7	70.9
M1	—	—	—	72.9

【0202】

ヘリカーゼタンパク質をコードするORF1ヌクレオチド配列をUS-1、US-2、M1及びP1単離物の各々の間で比較した。HEV US-1ゲノムのヘリカーゼタンパク質コーディング領域は配列番号89の残基2893～359

1で示されるのに対して、HEV US-2ゲノムのヘリカーゼタンパク質コーディング領域は配列番号164の残基2958～3656で示される。比較結果を表24に示す。

【0203】

【表25】

表24：ヘリカーゼ領域

	同一性 (%)			
	US-1	US-2	M1	P1
US-1	—	92.9	76.5	75.2
US-2	—	—	75.4	74.1
M1	—	—	—	76.2

【0204】

RNA依存性RNAポリメラーゼをコードするORF1ヌクレオチド配列をUS-1、US-2、M1及びP1単離物の各々の間で比較した。HEV US-1ゲノムのポリメラーゼコーディング領域は配列番号89の残基3634～5094で示されるのに対して、HEV US-2ゲノムのポリメラーゼコーディング領域は配列番号164の残基3699～5159で示される。比較結果を表25に示す。

【0205】

【表26】

表25：RNA依存性RNAポリメラーゼ領域

	同一性 (%)			
	US-1	US-2	M1	P1
US-1	—	93.1	72.9	75.3
US-2	—	—	73.6	75.8
M1	—	—	—	77.1

【0206】

更に、US-1、US-2、B1、B2、I2、C1、C2、C3、P1、C4及びI1株のORF1、ORF2及びORF3配列によりコードされるタンパク質のアミノ酸同一性/類似性をそれぞれ表26～28に示す。加えて、表27

及び28には、ブタ配列(S1)に対する比較も示す。表26～28において、類似性は表の右上半分に示し、同一性は表の左下半分に示した。

【0207】

【表27】

表26：ORF1の範囲のアミノ酸類似性／同一性

		類似性%											
同一性%		US-1	US-2	B1	B2	I2	C1	C2	C3	P1	C4	I1	M1
	US-1		97.8	86.0	85.7	84.4	85.9	86.2	84.9	86.4	85.7	86.3	85.4
	US-2	97.5		86.2	85.8	84.5	85.8	86.0	85.0	86.3	85.7	86.3	85.5
	B1	82.4	82.6		98.7	96.8	98.4	98.5	97.1	98.5	98.1	98.2	87.0
	B2	82.3	82.3	98.6		96.2	97.8	97.9	96.3	97.8	97.6	97.6	86.6
	I2	80.7	80.7	96.3	95.7		96.3	96.4	95.0	96.3	95.9	95.9	85.2
	C1	82.5	82.3	98.2	97.5	95.7		99.5	97.9	99.4	99.0	98.2	86.9
	C2	82.8	82.6	98.4	97.8	95.9	99.4		98.2	99.6	99.2	98.4	87.0
	C3	81.6	81.6	96.9	96.1	94.4	97.7	98.1		98.1	97.6	97.0	85.9
	P1	83.0	82.9	98.4	97.7	95.9	99.2	99.6	98.0		99.0	98.4	87.1
	C4	82.5	82.3	98.0	97.6	95.4	98.8	99.1	97.4	98.9		97.8	86.5
	I1	82.9	82.9	98.1	97.5	95.5	98.1	98.4	96.9	98.4	97.8		87.3
	M1	82.0	82.0	83.8	83.4	81.8	83.7	83.9	82.8	84.0	83.4	84.2	

【0208】

【表28】

表27：ORF2の範囲のアミノ酸類似性／同一性

		類似性 %													
同一性 %		US-1	US-2	B1	B2	I2	C1	C2	C3	P1	C4	I1	M1	S1	
	US-1		98.3	93.3	93.0	93.0	93.5	93.2	92.9	93.2	92.4	92.6	91.5	97.1	
	US-2	98.0		93.3	93.0	93.3	93.3	93.3	93.0	93.3	92.6	92.7	91.7	99.1	
	B1	91.8	91.8		98.9	99.1	99.8	99.2	99.2	99.5	98.8	98.9	94.8	93.0	
	B2	91.5	91.5	98.9		98.3	99.1	98.5	98.5	98.8	98.2	98.2	94.1	92.7	
	I2	91.5	91.8	99.1	98.3		99.2	98.9	98.6	99.2	98.5	98.6	94.5	91.5	
	C1	92.0	92.0	99.7	98.9	99.1		99.4	99.1	99.7	98.9	99.1	95.0	93.2	
	C2	91.7	92.0	99.1	98.3	98.8	99.4		98.8	99.4	98.6	98.8	94.7	93.0	
	C3	91.4	91.7	99.1	98.3	98.5	99.1	98.8		99.1	98.3	98.5	94.4	92.7	
	P1	91.7	92.0	99.4	98.6	99.1	99.7	99.4	99.1		98.9	99.1	95.0	93.0	
	C4	90.9	91.2	98.6	98.0	98.4	98.9	98.6	98.3	98.9		98.3	94.2	92.3	
	I1	91.1	91.4	98.5	97.7	98.2	98.8	98.5	98.2	98.8	98.0		94.7	92.4	
	M1	90.1	90.6	93.2	92.4	92.9	93.3	93.0	92.9	93.3	92.6	93.0		91.2	
	S1	97.7	98.9	91.7	91.4	91.9	91.8	91.7	91.4	91.7	90.9	91.1	90.2		

【0209】

【表29】

表28：ORF3の範囲のアミノ酸類似性／同一性

		類似性%												
		US-1	US-2	B1	B2	I2	C1	C2	C3	P1	C4	I1	M1	S1
同一性%	US-1		96.7	85.2	84.4	85.2	85.2	83.6	85.2	85.2	83.6	85.2	79.5	93.5
	US-2	96.7		85.2	84.4	85.2	85.2	83.6	83.6	85.2	83.6	85.2	81.1	96.7
	B1	84.4	84.4		98.4	100.0	100.0	98.4	98.4	100.0	98.4	98.4	87.0	83.7
	B2	83.6	83.6	98.4		98.4	98.4	96.7	96.7	98.4	96.7	96.7	87.0	82.9
	I2	84.4	84.4	100.0	98.4		100.0	98.4	98.4	100.0	98.4	98.4	87.0	83.7
	C1	84.4	84.4	100.0	98.4	100.0		98.4	98.4	100.0	98.4	98.4	87.0	83.7
	C2	82.8	82.8	98.4	96.7	98.4	98.4		96.7	98.4	97.6	96.7	85.4	82.1
	C3	84.4	82.8	98.4	96.7	98.4	98.4	96.7		98.4	96.7	96.7	85.4	82.1
	P1	84.4	84.4	100.0	98.4	100.0	100.0	98.4	98.4		98.4	98.4	87.0	83.7
	C4	82.8	82.8	98.4	96.7	98.4	98.4	97.6	96.7	98.4		96.7	85.4	82.1
	I1	84.4	84.4	98.4	96.7	98.4	98.4	96.7	96.7	98.4	96.7		88.6	83.7
	M1	78.7	80.3	87.0	87.0	87.0	85.4	85.4	87.0	85.4	88.6			79.7
	S1	93.5	96.7	82.9	82.1	82.9	82.9	81.3	81.3	82.9	81.3	82.9	78.9	

【0210】

加えて、メチルトランスフェラーゼタンパク質を規定するORF1アミノ酸配列をUS-1、US-2、M1及びP1単離物の各々の間で比較した。HEV US-1ゲノムによりコードされるメチルトランスフェラーゼタンパク質は配列番号91の残基1～231で示されるのに対して、HEV US-2ゲノムによりコードされるメチルトランスフェラーゼタンパク質は配列番号166の残基1～240で示される。比較結果を表29に示す。

【0211】

【表30】

表29：メチルトランスフェラーゼ領域

類似性(%)		同一性(%)			
		US-1	US-2	M1	P1
	US-1	—	98.7	91.3	88.7
	US-2	98.7	—	91.7	89.1
	M1	91.8	92.0	—	92.9
	P1	90.0	90.4	91.2	—

【0212】

プロテアーゼタンパク質を規定するORF1アミノ酸配列をUS-1、US-2、M1及びP1単離物の各々の間で比較した。HEV US-1ゲノムによりコードされるプロテアーゼタンパク質は配列番号91の残基424～697で示されるのに対して、HEV US-2ゲノムによりコードされるプロテアーゼタ

ンパク質は配列番号166の残基433～706で示される。比較結果を表30に示す。

【0213】

【表31】

表30：プロテアーゼ領域

類似性 (%)		同一性 (%)			
		US-1	US-2	M1	P1
	US-1	—	98.5	67.5	69.3
	US-2	97.8	—	67.1	68.6
	M1	73.3	73.3	—	76.6
	P1	74.4	74.0	72.2	—

【0214】

Yドメインタンパク質を規定するORF1アミノ酸配列をUS-1、US-2、M1及びP1単離物の各々の間で比較した。HEV US-1ゲノムによりコードされるYドメインタンパク質は配列番号91の残基207～424で示されるのに対して、HEV US-2ゲノムによりコードされるYドメインタンパク質は配列番号166の残基216～433で示される。比較結果を表31に示す。

【0215】

【表32】

表31：Yドメイン

類似性 (%)		同一性 (%)			
		US-1	US-2	M1	P1
	US-1	—	98.2	92.7	93.6
	US-2	98.2	—	92.7	93.6
	M1	94.0	94.0	—	93.1
	P1	94.5	94.5	91.7	—

【0216】

Xドメインタンパク質を規定するORF1アミノ酸配列をUS-1、US-2、M1及びP1単離物の各々の間で比較した。HEV US-1ゲノムによりコードされるXドメインタンパク質は配列番号91の残基789～947で示され

るのに対して、HEV US - 2ゲノムによりコードされるXドメインタンパク質は配列番号166の残基799～957で示される。比較結果を表32に示す。

【0217】

【表33】

表32：Xドメイン

類似性 (%)		同一性 (%)			
		US - 1	US - 2	M 1	P 1
	US - 1	—	97.5	82.4	80.5
	US - 2	97.5	—	81.8	79.9
	M 1	88.0	87.4	—	86.0
	P 1	84.3	83.6	83.0	—

【0218】

ヘリカーゼタンパク質を規定するORF1アミノ酸配列をUS - 1、US - 2、M 1及びP 1単離物の各々の間で比較した。HEV US - 1ゲノムによりコードされるヘリカーゼタンパク質は配列番号91の残基965～1197で示されるのに対して、HEV US - 2ゲノムによりコードされるヘリカーゼタンパク質は配列番号166の残基975～1207で示される。比較結果を表33に示す。

【0219】

【表34】

表33：ヘリカーゼ領域

類似性 (%)		同一性 (%)			
		US - 1	US - 2	M 1	P 1
	US - 1	—	99.1	89.7	91.0
	US - 2	99.1	—	90.6	91.8
	M 1	93.1	94.0	—	95.2
	P 1	94.0	94.8	91.0	—

【0220】

超可変領域を規定するORF1アミノ酸配列をUS - 1、US - 2、M 1及びP 1単離物の各々の間で比較した。HEV US - 1ゲノムによりコードされる

超可変領域は配列番号91の残基698～788で示されるのに対して、HEV

US-2ゲノムによりコードされる超可変領域は配列番号166の残基707～798で示される。比較結果を表34に示す。

【0221】

【表35】

表34：超可変領域

類似性 (%)		同一性 (%)			
		US-1	US-2	M1	P1
	US-1	—	82.4	25.0	27.7
	US-2	79.1	—	25.0	21.0
	M1	25.0	25.0	—	20.8
	P1	31.9	21.0	18.0	—

【0222】

RNA依存性RNAポリメラーゼを規定するORF1アミノ酸配列をUS-1、US-2、M1及びP1単離物の各々の間で比較した。HEV US-1ゲノムによりコードされるポリメラーゼは配列番号91の残基1212～1698で示されるのに対して、HEV US-2ゲノムによりコードされるポリメラーゼは配列番号166の残基1222～1708で示される。比較結果を表35に示す。

【0223】

【表36】

表35：RNA依存性RNAポリメラーゼ領域

類似性 (%)		同一性 (%)			
		US-1	US-2	M1	P1
	US-1	—	99.0	86.0	87.8
	US-2	99.0	—	86.2	87.7
	M1	89.7	89.9	—	92.6
	P1	91.6	91.6	89.5	—

【0224】

更に、本研究中HEV USタイプファミリーに属する幾つかの別のHEV単離物を同定した（以下の実施例13参照）。別の単離物をIt1（イタリア株）

、G1（第1ギリシャ株）及びG2（第2ギリシャ株）と称する。追加の配列比較を実施し、その中にIt1、G1及びG2配列を含め、その結果を下表36及び37に示す。表36は371塩基（123アミノ酸）ORF1断片の範囲のHEV単離物間のヌクレオチド及び推定アミノ酸同一性を示す。ORF1断片は配列番号89の残基26～396に対応する。表37は148塩基（49アミノ酸）ORF2断片の範囲のHEV単離物間のヌクレオチド及び推定アミノ酸同一性を示す。ORF2断片は配列番号89の残基3607～6454に対応する。表36及び37において示した単離物はビルマ（B1，B2）、中国（C1，C2，C3，C4）、インド（I1，I2）、パキスタン（P1）、メキシコ（M1）、ブタ（S1）、米国（US-1，US-2）、ギリシャ（G1，G2）及びイタリア（It1）である。

【0225】

HEV US-1及びHEV US-2のそれぞれのゲノムのヌクレオチド配列を表14で同定した他のHEV単離物の他のゲノムと一緒に用いた完全長ヌクレオチド配列の1対比較が好ましかった。比較の結果を表15に示す。ヌクレオチドレベルで、HEV US-1及びHEV US-2は相互に最も密接に関連しており、全ゲノムの範囲で92.0%の同一性であった。完全長ビルマ様単離物も92.0～98.8%の同一性を示した。US単離物はビルマ様単離物及びメキシコ単離物に対して73.5～74.5%同一であった。これは、75.0～76.1%というビルマ様単離物及びメキシコ単離物間のヌクレオチド同一性と似ている。これらのデータは、US単離物はビルマ株及びメキシコ株とは区別されるHEVの新規株変異体のメンバーであることを示す。

【0226】

各ゲノムの小部分、例えば個々のORFを分析するときに同様の同一性が見られる。これらの値をORF1、ORF2及びORF3のそれぞれについて表16～18に示す。各領域の範囲で、ビルマ単離物及びパキスタン単離物は93.1～98.9%の最高同一度を示す。メキシコ単離物はビルマ様単離物に対して73.6～90.1%の同一性であり、区別される。HEV US-1ヌクレオチド配列分析は、ビルマ様単離物及びメキシコ単離物に対する同一性が72%未満

であり、ORF 1 配列に対する相違が有意であることを示す。同様に、ORF 2 及びORF 3 配列の同一性はビルマ様単離物及びメキシコ単離物に対してそれぞれ79.1%及び86.9%未満であった。

【0227】

ヌクレオチドレベルで見られる変動は翻訳されたオープンリーディングフレームのアミノ酸の類似性及び同一性において反映される。ORF 1 は多分超可変領域が存在するために最も逸脱している産物である。US 単離物はこの領域の範囲で97.5%のアミノ酸同一性を有する(表26)。これは、ビルマ様ORF 1 タンパク質間で見られる94.4~99.6%同一性に類似している。US ORF 1 産物はビルマ様タンパク質及びメキシコタンパク質に対して80.7~83.0%同一である(表26)。これらの値は、ビルマ様単離物とメキシコ単離物間で見られる81.8~84.2%同一性に類似している。アミノ酸類似性の値は通常同一性の値よりも最高3.5%高く、これは保守的アミノ酸置換が多数あることを反映している。ORF 2 産物は、多分ウイルスクャプシドタンパク質としての役割のために最も保存される。US ORF 2 産物は相互に98.0%同一であり、ビルマ及びメキシコORF 2 タンパク質と90.1~92%同一であった(表27)。また、これらの領域はビルマ単離物間で見られる同一性(97.7~99.7%)を反映している。ビルマ単離物とメキシコ単離物間の同一性は、US 変異体と他の変異体間(92.4~93.3%)よりもわずかに高い。ORF 2 の範囲のアミノ酸類似性は同一性の値より約1.5%高い。HEV US-1 及びHEV US-2 のORF 3 産物は96.7%のアミノ酸同一性を有していた。ビルマ単離物は96.7~100%のアミノ酸同一性を示した。US 単離物のビルマ単離物及びメキシコ単離物に対するORF 3 アミノ酸同一性は、ビルマ単離物とメキシコ単離物間で見られる同一性(85.4~88.6%)よりも僅かに低く、78.7~84.4%であった(表28)。ORF 3 の範囲のアミノ酸類似性は通常同一性の値と同じであるが、幾つかの比較では類似性の値は同一性の値よりも1.0%未満しか高くないことが示された。これらのアミノ酸類似性及び同一性の値は、短アミノ酸配列の分析は完全長及び部分ヌクレオチド分析に対して同様の結果を与えることを示し、US 単離物がHEVの既に特

性付けられている単離物と密接に関連しており、遺伝的に区別されることを示唆している。

【0228】

表27及び28には、ブタで最近同定されたHEV様単離物(Mengら, *Proc. Natl. Acad. Sci. USA*, 94:9860-9865(1997))との1対アミノ酸配列比較が含まれている。ORF2/3領域の範囲の2021bpのみが特性付けられた(GenBank受託番号AF011921)。USブタ配列はヌクレオチドレベルでHEV US-1の対応領域と92%同一である。HEV US-1はアミノ酸レベルで最近同定されたブタウイルスに非常に類似していることに注目されたい。例えば、HEV US-1及びブタ株はそれぞれORF2及びORF3配列の範囲で97.1%及び93.5%の同一性を示す(表27及び28)。

【0229】

G9及びG20(それぞれGenbank受託番号X87306及びX87307)と呼ぶ中国からの2つのHEV単離物由来の210ヌクレオチドの部分配列が最近文献(Huangら, *J. Med. Virology*, 47:303-308(1995))に記載された。これらの断片は配列番号89の残基4533~4742に相同のヌクレオチド配列を表す。このコード化アミノ酸配列(69アミノ酸残基の長さ)は配列番号91の残基1512~1580と相同である。ヌクレオチド配列及び前記配列の推定されるアミノ酸配列の1対比較の結果を表38及び39に示す。結果は、G9及びG20単離物はこの領域の範囲のヌクレオチドレベルで相互に89%同一であることを示す。密接に関連しているビルマ単離物及びパキスタン単離物はこの領域の範囲で92.9%同一である。US-1単離物はこの領域の範囲で77.1%及び81.0%を示し、US-1単離物がこれらの単離物とはユニークであることが示唆される。G9及びG20配列はヌクレオチドレベルで最も密接に関連しているが、G20の推定アミノ酸翻訳はUS-1単離物からのUS配列と最も類似/同一である(表38)。これは、多分析に使用したアミノ酸の長さが短いためである可能性が高い。

【0230】

【表37】

表38：ORF1の210ヌクレオチドの範囲の同一性

	P a k	M e x	U S-1	G 2 0	G 9
B u r	92.9	74.8	75.7	78.1	76.7
P a k		75.2	76.7	78.1	76.7
M e x			77.1	75.2	71.9
U s-1				81.0	77.1
G 2 0					89.0

【0231】

【表38】

表39：ORF1の69アミノ酸の範囲の類似性／同一性

	P a k	M e x	U S-1	G 2 0	G 9
B u r	98.6/98.6	92.8/88.4	92.8/85.5	92.8/88.4	82.6/79.7
P a k		94.2/89.9	91.3/84.1	91.3/87.0	84.1/81.2
M e x			89.9/87.0	89.9/87.0	81.2/78.3
U s-1				100/95.7	88.4/88.1
G 2 0					88.4/87.0

【0232】

実施例7：系統発生的分析

HEVの新規なUSタイプ単離物及び他の単離物間の系統発生的関係を調べるためにヌクレオチド及びアミノ酸配列のアラインメントを実施した。前記アラインメントは、Wisconsin Sequence Analysis Package (ウィスコンシン州マジソンに所在のGenetic Computer Groups) バージョン9のプログラムPILEUPを用いて作成した。配列間の進化距離を、DNADISTプログラム (Kimura 2 - パラメーター法) を利用し、2.0のトランジション - トランスバージョン比及びPHYLIIPパッケージ (Felsenstein 1993, シアトルに所在のワシントン大学の遺伝学部門) バージョン3.5cのPROTDIST (Dayhoff PAMマトリックス) プログラムを用いて調べた。プログラムFITCH (Fitch - Margoliash法) を用いて系統樹を構築するためにコンピュータ化した距離を使用した。樹のロバストを、マルチプル配列アラインメ

ント(100組または1000組)をプログラムSEQBOOT、DNADIST、プログラムNEIGHBORの隣接結合方法及びCONSENSE(PHYLIPパッケージ)を用いてブーツストラップリサンプリングすることにより調べた。70%未満のブーツストラップ値は、系統発生的グループ化のための証拠を与えないと見做される(Muerhoffら, Journal of Virology, 71:6501-6508(1997))。最終樹をRETREE(PHYLIP)を利用して中点ルーティングオプションで作成し、グラフ出力をTREEVIEW(Computer Applied Biosciences, 12:357-358(1996))を用いて作成し、その結果を図5、6、10及び12に示す。

【0233】

完全ゲノムを用いる系統発生的分析

HEV US-1、HEV US-2及びHEVの他の公知の単離物間の関連度を広範囲に調べるために、ヌクレオチドアラインメントを実施した。完全長HEV US-1及びHEV US-2ゲノムを、完全ゲノムが入手可能な10個のHEVの他の単離物とアラインした(表14)。

【0234】

DNADISTプログラム(Kimura 2-パラメーター法)を利用して2.0のトランジション-トランシバージョン比で測定したHEV US単離物及びHEVの他の単離物のアラインメントに基づく系統発生的距離の調査から、US由来の単離物と他の地域由来の単離物の間にかなりの進化距離があることが分かる(表40)。計算した距離はアジアに由来する単離物間の密接した関係をも示す。ビルマ様グループ内では、完全長アラインメントから計算した最長距離は塩基あたり0.0850ヌクレオチド置換である。このグループのメンバーとUS単離物間の最短距離は0.3322置換である。メキシコ株は、ビルマ様グループに対して0.3055~0.3132置換、US単離物に対して0.3322~0.3462置換という類似の距離を示す。HEV US-1とHEV US-2間の遺伝的距離は0.0812置換であり、これはビルマ様単離物間で見られる距離に類似している。分析したウイルス配列間の相対的進化距離は、図

5に示す根無し系統樹を精査することにより明らかである。ここで、枝の長さは進化距離に比例する。系統樹において、ビルマ様単離物、メキシコ単離物及び米国単離物はそれぞれ主要枝を表す。更に、原型ウイルスの枝は、ブーツストラップ値が100%であることから裏付けられる。ゲノムの小セグメント（例えば、ORF1、ORF2またはORF3）を個別に分析すると、完全長配列で得た樹に似た樹が生じ、図5に示す。これらの分析から、HEV US単離物がHEVの異なる株または変異体を表し、最も異なるビルマ様単離物のようにHEV US-1及びHEV US-2は相互に類似していることが分かる。

【0235】

【表39】

表40：完全長配列の範囲の系統発生的距離

	B1	B2	C1	C2	C3	C4	I1	I2	P1	M1	US-1
B1											
B2	0.0149										
C1	0.0643	0.0697									
C2	0.0680	0.0733	0.0136								
C3	0.0663	0.0734	0.0178	0.0132							
C4	0.0574	0.0611	0.0304	0.0290	0.0329						
I1	0.0677	0.0728	0.0645	0.0625	0.0647	0.0681					
I2	0.0403	0.0477	0.0820	0.0849	0.0846	0.0776	0.0832				
P1	0.0693	0.0751	0.0178	0.0120	0.0172	0.0335	0.0633	0.0850			
M1	0.3096	0.3120	0.3086	0.3089	0.3091	0.3132	0.3120	0.3259	0.3055		
US-1	0.3406	0.3418	0.3360	0.3345	0.3367	0.3445	0.3322	0.3464	0.3363	0.3462	
US-2	0.3413	0.3408	0.3370	0.3361	0.3374	0.3445	0.3333	0.3461	0.3377	0.3367	0.0812

【0236】

ブタHEV由来のORF2/ORF3に対する比較

最近発表されたブタHEV及びヒトHEV US-1及びHEV US-2単離物間の関連を調べるために、完全ORF2及びORF3の範囲のヌクレオチド配列を上で使用した10個の完全長配列からの類似領域を用いて比較した（表14）。系統発生的分析により、US及びブタHEV単離物間で1位置あたり0.0799～0.0810ヌクレオチド置換の遺伝的距離が生ずる（表41）。この値は、最も遠いビルマ様単離物間で見られた値と似ている。ORF2/3ヌ

クレオチド配列を分析したときに根無し系統樹でUS及びブタ単離物グループは近い(図6参照)。これらの単離物は、メキシコ単離物及びビルマ様単離物から離れた系統発生的グループを形成する。これらのグループ分けは、ブーティストラップ値が100%であることから裏付けられる。

【0237】

【表40】

表41：米国ブタ及びヒトHEV単離株間の系統発生的距離

	US-2	USブタ	ビルマ	メキシコ
US-1	0.0799	0.0810	0.2441-0.2495	0.2671
US-2		0.0795	0.2409-0.2479	0.2486
米国ブタ			0.2348-0.2485	0.2615
ビルマ			0.0119-0.0716	0.2183-0.2248

【0238】

実施例8：HEV血清学的研究

(A) 背景

早期研究で、HEV感染の診断に有用なエピトープがHEVのビルマ株及びメキシコ株の両方のORF2及びORF3のカルボキシル末端近くにあることが分かっている。以後M3-2及びM4-2と称するメキシコ株由来の2つの抗原はそれぞれORF2及びORF3のカルボキシル末端近くに42及び32アミノ酸を含む(Yarboroughら, Journal of Virology, 65: 5790-5879(1991))。以後B3-2及びB4-2タンパク質と称するHEVのビルマ株由来の2つの抗原はそれぞれORF2及びORF3のカルボキシル末端近くに42及び32アミノ酸を含む(上掲したYarboroughら(1991))。HEVに対するIgG、IgA及びIgMクラス抗体を検出するように設計された診断テストが上記抗原領域に基づいて開発された。HEVに対する抗体の検出を潜在的に高めるために、完全長ORF3を含む別のHEV組み換えタンパク質(Dawsonら, Journal of Virology Methods, 38: 175-186(1992))またはORF2タンパク質由来の別のアミノ酸配列(上掲したDawsonら(1993))を作

成した。比較研究で、元の組み換えタンパク質及び合成ペプチド（B4-2、B3-2、M3-2、M4-2）が公知の急性HEV感染症例においてHEVに対する抗体を検出する際により大きな組み換えタンパク質と同程度に有効であったことが分かっている。HEVに対する抗体を検出するための認可テストがAbbott Laboratoriesで製造されており、完全長のビルマ株ORF3タンパク質及びビルマ株ORF2タンパク質のカルボキシル327アミノ酸から構成されている。

【0239】

初期の血清学的研究でB3-2、B4-2、M3-2及びM4-2の有用性を立証した後、M3-2及びB3-2抗原ペプチドの一部を形成しないHEVのビルマ及びメキシコ株の両方のORF2のカルボキシ末端に更に6個のアミノ酸残基があることが確認された。ORF2及びORF3のカルボキシ末端はHEVのビルマ株及びメキシコ株にとって重要であることが判明したので、ゲノムの前記領域に対応する合成ペプチドをHEVのUS-1株に対して作成した。ORF2のカルボキシ末端の48アミノ酸に対応する合成ペプチドをHEVのビルマ株及びメキシコ株（それぞれ、配列番号172及び170）に対して作成し、B3-2e及びM3-2e（ここで、“e”は延長アミノ酸配列を指す）と称する。更に、HEV US-1 ORF3のカルボキシル末端の33アミノ酸を表す合成ペプチドをHEVのビルマ株及びメキシコ株（それぞれ、配列番号171及び169）に対して作成し、B4-2及びM4-2と称する。HEV US-1株のORF2（配列番号174）内のエピトープに基づく合成ペプチドをUS3-2eと称する。HEV US-1 ORF3（配列番号173）のカルボキシ末端のエピトープに基づく合成ペプチドをUS4-2eと称する。HEVのメキシコ株及びビルマ株から誘導される前記ペプチドの各々を合成し、固相上にコートし、合成ペプチドの相対的有用性を調べるためにELISA試験で用いた。

【0240】

表42に示すように、HEV US-1とHEVのビルマ株、メキシコ株及びパキスタン株間のアミノ酸同一性は、ORF2内の3-2eエピトープを含むアミノ酸に対して約87.5～約91.7%、ORF3内の4-2エピトープを含

むアミノ酸に対して約63.6～約72.7%である。理論に束縛されるものはないが、エピトープをコードする領域が変化するならば前記ウイルスに対する株特異性抗体応答がありそうである。

【0241】

【表41】

表42：類似性／同一性

	3-2eペプチド			4-2ペプチド		
	Pak	Mex	US-1	Pak	Mex	US-1
Bur	100/97.9	91.7/91.7	93.7/91.7	100/100	72.7/72.7	72.7/72.7
Pak		91.7/91.7	93.7/91.7		72.7/72.7	72.7/72.7
Mex			89.6/87.5			63.6/63.6

【0242】

(B) 急性HEV感染の診断におけるELISAの使用

ヒトにおける急性HEV感染の多くの症例は1つ以上のHEV組み換えタンパク質または合成ペプチドに結合するIgMクラス抗体を有していることが報告されている。ヒトがHEVに対するIgMクラス抗体を有していない場合には、急性HEV感染の診断の基準は血清学に基づいて行うことができず、RT-PCR及び/または病因物質としてのHEVを証明するための他の検査が必要となることがある。

【0243】

(C) 合成ペプチドの作成

ペプチドを、Rainin Symphony Multiple Peptide SynthesizerにおいてN-メチルモルホリンにより与えられるin situ活性化による(HBTU)カップリング化学を用い、各残基で45分のカップリング時間及び所定の残基で2倍カップリングとして、0.025マイクロモルスケールで標準FMOC固相ペプチド合成により製造した。樹脂を通常通り開裂すると非保護ペプチドが得られ、その後にエーテル沈降及び洗浄を行った。合成したペプチドを表43に示す。

【0244】

【表42】

表 4 3

ペプチド	配 列	配列番号
B 3-2e	TLDYPARAHTFDDFCPECRPLGLQGCAFQSTVAELQRLKMKVVGKTREL	配列番号 :172
B 4-2	ANPPDHSAPLGVTRPSAPPLPHVVDLPQLGPRR	配列番号 :171
M 3-2e	TFDYPGRAHTFDDFCPECRALGLQGCAFQSTVAELQRLKVKVVGKTREL	配列番号 :170
M 4-2	ANQPGHLAPLGEIRPSAPPLPPVADLPQPLRR	配列番号 :169
US 3-2e	TVDYPARAHTFDDFCPECRTLGVGCAFQSTIAEVQRLKMKVVGKTREV	配列番号 :174
US 4-2	DSRPAPSVPLGVTSAPPLPPVVDLPQLGLRC	配列番号 :173

【0245】

(D) 合成ペプチドの分析

合成したペプチドを、そのアミノ酸組成について分析した。小規模合成（0.025マイクロモル）で製造した粗なペプチドの品質を各溶媒に0.1%（v/v）2-トリフルオロ酢酸（TFA）を添加したアセトニトリル/水勾配を用いてC18逆相高速液体クロマトグラフィーにより分析した。分析クロマトグラムから、各合成からの主要ピークを集め、溶離液を質量分析法（エレクトロスプレー及び/またはレーザー脱着質量分析法）により分析した。ペプチド（小規模及び/または大規模）は、各溶媒に0.1% TFAを添加したアセトニトリル/水勾配を用いてC18逆相HPLCを用いて精製した。主要ピークを集め、使用するまで凍結乾燥した。

【0246】

(E) ELISA試験

HEV US-1エピトープの有用性を、1/4インチのポリスチレンビーズを各ペプチドでコートすることにより調べた。具体的には、ペプチドを水または水+氷酢酸において可溶化し、リン酸緩衝液（pH7.4）中に10μg/mlを含むように希釈した。全部で60個のポリスチレンビーズを10μg/mlのペプチド溶液（14ml）と一緒にシンチレーション管に添加し、リン酸緩衝食塩液（PBS）において56℃で2時間インキュベートした。インキュベート後、液体を吸引し、0.1% トリトンX-100（登録商標）を含有する緩衝液と交換した。ビーズをこの溶液に60分間さらし、流体を吸引し、ビーズをPBS緩衝液で2回洗浄した。次いで、ビーズを5% ウシ血清アルブミン溶液と4

0 で60分間インキュベートした。インキュベート後、流体を吸引し、ビーズをPBSで濯いだ。生じたビーズを5% スクロース含有PBS中に30分間浸した。次いで、流体を吸引し、ビーズを風乾した。

【0247】

1つの研究で、1/4インチのポリスチレンビーズを各種濃度の合成ペプチド（1ロットあたり約50個のビーズ）でコートし、ネガティブコントロールとして抗-HEVセロネガティブヒトからの血清、ポジティブコントロールとしてHEV患者からの回復期の血清を用いて（以下に記載する）ELISA試験で評価した。最高のポジティブコントロール信号/ネガティブコントロール信号比を与えるビーズコーティング条件をビーズコーティング方法を大規模とするために選択した。HEV US-1 ORF2及びORF3に対して2つの1,000ビーズロットを作成した後、以下のように使用した。

【0248】

血清または血漿サンプルをサンプル希釈剤で希釈し、サンプル中の抗体が固定化抗原に結合し得る条件で抗原をコートした固相と混合した。洗浄後、生じたビーズを固相に結合させたタマリンまたはヒト抗体に結合するホースラディッシュペルオキシダーゼ（HRPO）標識した抗-ヒト抗体と混合した。カットオフ値以上の信号を発生するサンプルを反応性で見做した。

【0249】

より具体的には、好ましいELISAフォーマットでは、抗原コートした固相を予めサンプル希釈剤（動物血清及びノニオン性洗剤を含有する緩衝溶液）で希釈した血清と接触させなければならない。具体的には、血清（10 μ l）をサンプル希釈剤（150 μ l）で希釈し、攪拌した。次いで、予備希釈したサンプル（10 μ l）をELISAプレートの各ウェルに添加した後、サンプル希釈剤（200 μ l）及び抗原コートしたポリスチレンビーズを添加した。次いで、ELISAプレートをDynamic Incubator（Abbott Laboratories）で常に攪拌しながら室温で1時間インキュベートした。インキュベート後、流体を吸引し、ウェルを蒸留水（1洗浄あたり5ml）で3回洗浄した。次いで、上記したようにコンジュゲート希釈剤（動物血清及びノニオン

性洗剤を含有する緩衝溶液) 中で希釈したHRPO標識ヤギ抗-ヒト免疫グロブリン(200 µl)を各ウェルに添加し、ELISAプレートを1時間インキュベートした。次いで、ウェルを蒸留水で3回洗浄して抗原及び結合免疫グロブリンを含有するビーズを各ウェルから除去した後、0.1 M クエン酸塩緩衝液(pH 5.5)、0.3% o-フェニレンジアミン-2HCl及び0.02%過酸化水素を含む溶液(300 µl)と共に試験管に入れた。室温で30分後、1 N 硫酸を添加して反応を停止した。492 nmでの吸光度を記録した。生じた色の濃度は試験サンプル中に存在する抗体の量に直接比例した。各群のサンプルに関して、多分HEVエピトープに対する抗体を含んでいるサンプルを該抗体を含んでいないサンプルから分離するように予備カットオフ値を設定した。

【0250】

パネル1：プレスクリーニングしたパネルの検査

HEV US-1株由来のエピトープの有用性を立証するために、試料パネルをHEV US-1アミノ酸配列に基づくELISAにより検査した(表44)。これらのサンプルについては、上記し、公表文献(上掲したDawsonら(1993), 上掲したPaul(1993))に記載されているように既存のペプチド及び認可されている抗-HEV(Abbott Laboratories)の組合せを用いてHEVに対する抗体についてプレスクリーニングした。

【0251】

前記パネルの最初の10個は、HEVのビルマ株及びメキシコ株由来の組み換えタンパク質及びペプチドの組合せを用いる分析により血清がHEVに対する抗体に関して陰性であった米国の志願血液ドナーから採取した試料であった。すべての試料がHEV US-1由来のELISAで非反応性であった。5個の別の試料を、急性肝炎を患っており、血清がHEV組み換え抗原及びHEVのビルマ株及びメキシコ株に基づく合成ペプチドに対するIgG及びIgMクラス抗体に対して反応性であったので急性HEV感染と診断された患者から得た。5個の試料のうち3個はエジプトから入手し、1個はインド、1個はノルウェー(旅行者)から入手した。HEV RNAは上記患者のうち5人すべてでRT-PCRにより検出された。5個の試料をHEV US-1単離物に対する抗体について検

査し、I g G及びI g Mの両クラスの抗体を各症例で検出した(表44)。よって、前記データから、H E VのU S - 1株由来の合成ペプチドがH E Vへの接触を診断し、急性H E V感染を診断するのに有用であることが裏付けられる。

【0252】

【表43】

表44

検査試料	認可 抗-H E V		米国単離株			
			I g G		I g M	
検査項目	IgG	IgM	4-2	3-2e	4-2	3-2e
ネガティブ コントロール	0.061	0.084	0.031	0.041	0.071	0.109
ポジティブ コントロール	0.567	1.051	1.606	1.619	1.376	1.798
米国志願 ドナー						
T G 8 2 7	—	—	—	—	—	—
E G 5 4 9	—	—	—	—	—	—
E C 7 6 0	—	—	—	—	—	—
R F 7 6 2	—	—	—	—	—	—
R F 7 6 2	—	—	—	—	—	—
R G 7 3 0	—	—	—	—	—	—
N H 7 7 0	—	—	—	—	—	—
A S 7 0 5	—	—	—	—	—	—
B W 4 9 4	—	—	—	—	—	—
C D 6 4 8	—	—	—	—	—	—
エジプト						
7	+	+	+	+	+	+
9	+	+	+	+	+	+
1 2	+	+	+	—	+	+
インド						
5 4 3	+	+	+	+	+	+
ノルウエー						
M 1	+	+	+	+	+	+

【0253】

パネル2：H E V U S - 1単離物の生物学的ソース中のH E Vに対する抗体
の検出

一連の血液を実施例1に記載の患者から入手し、その血清をHEV US-1株に対する生物学的ソースとして使用した。HEVのビルマ株及びメキシコ株で得た血清学的データに基づいて、前記患者はHEVに対する検出可能なIgMクラス抗体がなかったのでHEV陰性と誤診されたであろう。しかしながら、HEV US-1株に対するIgMクラス(表45)及びIgGクラス(表46)抗体が4個すべての採血日で検出された(表45及び46)。この患者の血清をHEV US-3-2e及びUS-4-2ペプチドに対するIgM及びIgGクラス抗体の存在について分析したならば、急性HEV感染陽性と診断されたであろう。この診断は、患者が急性肝炎を患っており、最も重要なことは血清サンプル中にHEV US-1株RNAが検出され得たという所見から裏付けられる。これらのデータは、HEV US-1株由来の合成ペプチドがHEVのために急性肝炎とより正確に診断するのに有用であり得ることを示している。

【0254】

【表44】

表45

試料	I g M : O R F 3 合成ペプチド4-2 単離物			I g M : O R F 2 合成ペプチド3-2 e 単離物		
	ビルマ	メキシコ	US-1	ビルマ	メキシコ	US-1
ネガティブコントロール	0.059	0.081	0.031	0.142	0.065	0.109
ポジティブコントロール	0.854	0.985	1.363	1.309	0.579	1.798
USP-1						
入院後8日目	-	-	+	-	-	+
入院後9日目	-	-	+	-	-	+
入院後10日目	-	-	+	-	-	+
入院後37日目	-	-	+	-	-	+

【0255】

【表45】

表 4 6

試料	I g G : ORF3合成ペプチド4-2 単離物			I g G : ORF2合成ペプチド3-2e 単離物		
	ビルマ	メキシコ	US-1	ビルマ	メキシコ	US-1
ネガティブコントロール	0.039	0.055	0.031	0.034	0.057	0.041
ポジティブコントロール	1.296	0.666	0.941	1.322	0.893	1.041
USP-1						
入院後8日目	-	-	+	-	-	+
入院後9日目	-	-	+	-	-	+
入院後10日目	-	-	+	-	-	+
入院後37日目	-	-	+	-	-	+

【0256】

パネル3：潜在的な急性HEV感染の他の症例

ビルマ株及びメキシコ株に対するI g Mクラス抗体が陰性であった急性肝炎と診断された50人の患者から得た血清を集めた。50個の血清サンプル中10個

はH E VのU S株に対する抗体について陽性であった(表47及び48)。上記サンプルに対してR T - P C Rを実施したが、10個のサンプルはいずれもH E V R N A陽性ではなかった。よって、本実施例で立証されているように、患者の血清をH E V U S - 1に対する抗体の存在について分析したとき不顕性ウイルス肝炎は急性H E V感染と診断され得る。

【0257】

【表46】

表47

試料	I g M : ORF3合成ペプチド4-2 単離物			I g M : ORF2合成ペプチド3-2 e 単離物		
	ビルマ	メキシコ	US-1	ビルマ	メキシコ	US-1
ネガティブコントロール	0.059	0.081	0.031	0.142	0.065	0.109
ポジティブコントロール	0.854	0.985	1.363	1.309	0.579	1.798
US	-	-	-	-	-	+
急性非A-E	-	-	-	-	-	+
SH755	-	-	-	-	-	+
DT314	-	-	-	-	-	+
EH673	-	-	-	-	-	+
SG560	-	-	-	-	-	+
SG681	-	-	-	-	-	-
N11C10	-	-	+	-	-	+
35	-	-	+	-	-	+
52	-	-	-	-	-	+
161	-	-	-	-	-	+
175	-	-	-	-	-	+

【0258】

【表47】

表48

試料	IgG: ORF3合成ペプチド4-2 単離物			IgG: ORF2合成ペプチド3-2e 単離物		
	ビルマ	メキシコ	US-1	ビルマ	メキシコ	US-1
ネガティブコントロール	0.039	0.055	0.031	0.034	0.057	0.041
ポジティブコントロール	1.296	0.666	0.941	1.322	0.893	1.041
US	-	-	-	-	-	-
急性非A-E	-	-	-	-	-	-
SH755	-	-	-	-	-	-
DT314	-	-	-	-	-	-
EH673	-	-	-	-	-	-
SG560	-	-	-	-	-	-
SG681	-	-	-	-	-	+
N11C10	-	-	-	-	-	-
35	-	-	-	-	-	+
52	-	-	-	-	-	-
161	-	-	-	-	-	-
175	-	-	-	-	-	-

【0259】

実施例9：動物伝染研究

マカクザル (*Macaca fascicularis*) はテキサス州サンアントニオに所在の Southwest Foundation for Bio

medical Research (SFB R) から入手した。動物は、ヒトに対するケア及び霊長類の倫理的使用を保証すべく SFB R が確立したガイドラインに従って維持し、モニターした。血清 ALT のベースラインレベルを得るために接種前少なくとも 4 週間の間週 2 回血清を採取した。ベースラインの平均 + 標準偏差の 3.5 倍に基づいてカットオフ (CO) 値を得た。2 匹のマカクに HEV 陽性 USP - 1 血清 (0.4 ~ 0.625 ml)、1 匹のマカクには HEV 陽性 USP - 2 血清 (2.0 ml) を静注により接種した。血清及び糞サンプルを接種後 (PI) 最長 6 週間まで週 2 回集めた。血清を ALT の変化について検査し、CO 以上の値は陽性で見做し、肝障害が示唆された。血清試料を実施例 8 に上記されているように HEV に対する抗体について検査した (表 49、図 7)。血清及び糞試料を RT - PCR により HEV RNA について検査した。マカク血清 (25 ~ 100 µl) を QIAamp Viral RNA キット (Qiagen) を用いて抽出した。10% 糞懸濁液を実施例 1 に記載されているように抽出した。RT - PCR を実施例 12 に下記するように実施した (図 7)。

【0260】

USP - 1 血清 (0.4 ~ 0.625 ml) を 2 匹のマカクザルに静脈内接種しても感染は生じなかったが (データ示さず)、患者 US - 2 の血清 (2.0 ml) を接種するとウイルス血症が生じ、血清中の肝酵素レベルが上昇した (図 7)。HEV RNA は PI 15 日目に糞試料中でまず検出され、PI 64 日目まで陽性のままであった。PI 28 ~ 56 日の間に集めた血清試料は HEV RNA 陽性であった。ALT 値の上昇が PI 15 日目、44 ~ 58 日目、72 日目及び 93 日目に記録され、PI 51 日目がピーク値 (116 IU/L) であった。

【0261】

抗体応答を評価するために 4 - 2 及び 3 - 2 e ペプチドに対してビルマ、メキシコ及び米国配列に基づく 6 つの ELISA を利用した。US 3 - 2 e ペプチドアッセイのみで測定可能な応答が認められ (表 49)、ビルマまたはメキシコペプチドに対する交差反応性は見られなかった。HEV に対する IgM クラス抗体は PI 28 ~ 58 日目に検出された。この後に、PI 44 日目に強い抗 - HEV - IgG 応答が見られた。

【 0 2 6 2 】

【表 4 8】

表 4 9

年月日	DPI	ALT	AST	GGT	IgG S/N
06/04/97	-82	35	37	102	1.4
06/06/97	-80	39	32	90	
06/11/97	-75	38	36	100	
06/13/97	-73	36	46	86	
06/18/97	-68	45	30	85	
06/20/97	-66	43	37	87	
06/25/97	-61	37	30	92	
06/27/97	-59	42	36	87	
08/25/97	0	41	36	107	1
08/27/97	2				
09/02/97	8	34	34	102	
09/04/97	10	34	31	91	
09/09/97	15	58	42	108	0.8
09/10/97	16	44	45	93	
09/15/97	21	35	32	86	
09/17/97	23	49	71	88	
09/22/97	28	39	33	86	1.2
09/24/97	30	40	37	90	
09/29/97	35	41	40	80	
10/01/97	37	48	58	90	1.1
10/03/97	39				
10/06/97	42	45	33	89	
10/08/97	44	58	38	94	6.2
10/15/97	51	116	62	89	11.9
10/20/97	56	87	38	83	33.6
10/22/97	58	76	43	85	29.9
10/28/97	64	45	42	88	17.2
10/29/97	65	46	34	88	
11/03/97	70	39	54	85	
11/05/97	72	54	47	88	13.3
11/10/97	77	47	33	93	
11/12/97	79	50	38	93	12.4
11/17/97	84	46	31	91	10.4
11/19/97	86	52	41	88	
11/26/97	93	67	104	109	7.2
12/03/97	100	36	36	108	
12/09/97	106	38	34	115	
12/10/97	107	36	29	103	2.1

【 0 2 6 3 】

実施例 10：組み換えタンパク質 ELISA

(A) 組み換え構築物

HEV US ゲノムの ORF 2 及び ORF 3 由来の HEV US 配列によりコードされる大腸菌由来組み換えタンパク質を、pJOorf3-29 (配列番号)、cksoorf2m-2 (配列番号 192) 及び CKS ORF32M-3 (配列番号 193) と称される CMS-KDO シンターゼ (CKS) との融合タンパク質として、或いは p1orf3-12 (配列番号 194)、p1orf2-2.6 (配列番号 195) 及び PLORF-32M-14-5 (配列番号 196)

と称される非融合タンパク質として発現させた。組み換え融合タンパク質の構築には、米国特許5,124,255号明細書に記載されているクローニングベクターpJO201を用いた。このベクターを制限エンドヌクレアーゼEcoRI及びBamHIで消化して、CKSとのインフレームでHEV-US配列をクローニングさせた。組み換え非融合タンパク質の構築には、ラムダpL発現ベクターpKRR826を用いた。このベクターを制限エンドヌクレアーゼEcoRI及びBamHIで消化して、HEV-US配列をリボソーム結合サイトの直ぐ下流でクローニングさせた。ベクター系が強力なラムダプロモーターを含んでいるので、温度を30 から温度感受性リプレッサータンパク質を不活化する42にシフトすることにより異種タンパク質合成を誘導する。構築物をクローン化し、前記HEVタンパク質を発現させるために大腸菌K12株HS36細胞に形質転換させた。

【0264】

HEV-US配列をHEV-US-2ヒト血清またはマカク13906糞試料から抽出した核酸から増幅し、実施例5に上記したように逆転写した。ORF2のカルボキシル半分を含む(すなわち、配列番号167のアミノ酸残基番号334~660をコードする)ORF2配列を、EcoRI制限サイト及びATG開始コドンを含むセンスプライマー(配列番号208)とFLAG(Eastman-Kodak)と称されるユニークなペプチド配列、2つの逐次TAA終止コドン及びBamHI制限サイトを含むアンチセンスプライマー(配列番号198)を用いて作成した。PCR反応(50μl)を製造業者が推奨するようにLA-TAQ(Takara)試薬を用いて始めた。サイクリング条件には、94×20秒、55×30秒、72×2分の40サイクルを含めた。増幅の前に94×1分、その後72×10分を含めた。産物をEcoRI及びBamHIで消化し、所望のベクターに連結させた。制限サイト間のCKS融合クローンのヌクレオチド配列を配列番号192に示し、その翻訳物を配列番号199に示す。制限サイト間の非融合クローンのヌクレオチド配列を配列番号195に示し、その翻訳物を配列番号200に示す。全ORF3(アミノ酸1~122)を包含するORF3配列を、EcoRI制限サイト及びATG開始コドンを含むセ

ンスプライマー（配列番号201）とFLAGと称されるユニークなペプチド配列、2つの逐次TAA終止コドン及びBamHI制限サイトを含むアンチセンスプライマー（配列番号202）を用いて作成した。PCR反応（50 μ l）を実施例5に記載のようにQiagen試薬を用いて開始した。サイクリング条件には、94 $\times$ 20秒、55 $\times$ 30秒、72 $\times$ 1分の35サイクルを含めた。増幅の前に94 $\times$ 1分のインキュベーション、その後に72 $\times$ 10分延長した。生じた産物をEcoRI及びBamHIで消化し、所望のベクターに連結させた。制限サイト間のCKS融合クローンのヌクレオチド配列を配列番号191に示し、その翻訳物を配列番号203に示す。制限サイト間の非融合構築物であるクローンを表すヌクレオチド配列を配列番号195に示し、その翻訳物を配列番号204に示す。

【0265】

更に、完全長ORF3（アミノ酸1～123）及びORF2（アミノ酸334～660）のカルボキシル半分を包含するキメラ構築物を作成した。PCR反応物（100 μ l）中の鋳型として、配列番号191及び配列番号192を含むプラスミド（約100mg）を用いた。PCR緩衝液及び酵素はLA-TAQキット（Takara）のものであり、製造業者の指示に従って用いた。ORF3を配列番号201及び205に示すプライマーを用いて増幅させた。配列番号205のアンチセンスプライマーは配列番号191のカルボキシル末端からFLAG配列及び停止コドンを排除し、ATG開始コドンを含まない配列番号192と同一の配列を含む。ORF2を配列番号208及び198に示すプライマーを用いて増幅させた。サイクリング条件はLA-TAQを用いて上記した通りであった。生じた産物を1.2%アガロースゲルで分画し、切除した。DNAを製造業者（Bio101）が記述しているようにGeneClean IIを用いてゲルスライスから単離した。産物をガラスビーズからH₂O（15 μ l）に溶出させた。ほぼ等モル比の各産物（ORF3産物10 μ l及びORF2産物1 μ l）を、1 $\times$ PCR緩衝液、dNTPs（0.5 μ l）及びLA-TAQ（Takara）（0.25 μ l）を用いて末端充填反応物（25 μ l）で混合した。この反応を以下のサイクルにかけた。94 $\times$ 1分；94 $\times$ 20秒、55 $\times$ 30秒

及び72 × 10分の10サイクル; 72 × 10分。この反応物(5 µl)を、LA T A Qキット(Takara)及び配列番号201及び198のプライマーを用いる増幅反応物(100 µl)に入れた。サイクリング条件は、94 × 1分、その後90 × 20秒、55 × 30秒及び72 × 1.5分の35サイクルであった。この後に、72 × 10分及び4 浸漬を設けた。適当なサイズの産物を制限酵素EcoRI及びBamHIで消化した。この産物をpJO201に連結し、同定されている適当な配列(配列番号193、その翻訳物を配列番号206に示す)を用いてクローン化した。生じた産物をpKRR826に連結し、同定されている適当な配列(配列番号196、その翻訳物を配列番号207に示す)を用いてクローン化した。

【0266】

(B) タンパク質発現及び精製

C K S構築物を2つの培養物(500 ml)(4時間誘導)において米国特許第5,312,737号明細書に記載されているように発現させた。P_L構築物は上記のように発現させた。誘導大腸菌培養物の凍結細胞ペレットをタンパク質精製のための出発物質として使用した。細胞をリゾチーム、DNase及びプロテイナーゼ阻害剤を含む緩衝液において溶解させた。可溶性タンパク質を11,000 × gで遠心して不溶性(封入体)タンパク質から分離した。組み換えタンパク質の溶解度をドデシル硫酸ナトリウム(SDS)ポリアクリルアミドゲル電気泳動(PAGE)及びFLAG(登録商標)M2を用いる抗体ウェスタンブロットから推定した。

【0267】

可溶性組み換えタンパク質を、適当な緩衝液に交換後FLAG(登録商標)M2抗体アフィニティーゲルを用いるアフィニティークロマトグラフィーにより精製した(Surowyら, Journal of General Virology, 78:1851-1859(1997))。所要により、サンプル及びクロマトグラフィー緩衝液に10 mM β-メルカプトエタノールを含有させたSephacryl(登録商標)S-200ゲル濾過クロマトグラフィーにかけて更に精製した。精製タンパク質を、280 nmでの吸光度を測定することによ

り定量化した。吸光度をタンパク質 1 mg に換算するために推定吸光係数 1 を用いた。タンパク質純度を、SDS PAGEにより分画化したタンパク質を予め純度を測定した標準物質を用いて走査デンストメトリー (Molecular Dynamics) にかけて測定した。

【0268】

(C) ELISA

組み換えHEV US構築物の潜在的有用性を調べるために、固相ELISAを実施し、評価した。すべての組み換えHEV USタンパク質を下記するように固相上にコートした。簡単に説明すると、100mM リン酸ナトリウム (pH 7.6) 中に希釈した0.5~10 µg/mlの濃度の量のPJORF3-29を用いて1/4" ポリスチレンビーズをコートした。1つの濃度条件あたり60個のビーズを緩衝液 (約14ml) においてコートし、40℃で2時間ひっくり返して回転させた。コーティング溶液を吸引し、コーティングの残りの手順は実施例8のセクションE、第1段落に記載されているように実施した。

【0269】

ELISAをPJOrf3-29コートビーズを用いて実施した。簡単に説明すると、血清または血漿を上記したSpecimen Diluent (SpD) で1:16に希釈した。次いで、この調製物の10 µlアリコート反応トレイのウェルに添加した後、SpD (200 µl) を添加した。1ウェルあたり1個のコートビードを添加し、Dynamic Incubator (Abbott Laboratories) を用いて動的モードにおいて37℃で1時間インキュベートした。インキュベート後、流体を吸引し、各ビードを脱イオン水で3回洗浄した (洗浄1回あたり5ml)。次いで、ビーズをHPRO標識したヤギ抗-ヒトIgGまたはIgMコンジュゲート (200 µl) とインキュベートし、上記したコンジュゲート希釈剤で希釈し、37℃で30分間インキュベートした。次いで、上記のようにコンジュゲートを吸引し、ビーズを洗浄した。発色及び吸光度を実施例8のセクションEに記載されているように測定した。

【0270】

前記構築物の免疫反応性を確認するために、実施例9に記載のHEV US-

2を実験的に感染させたマカク#13903からの一連の血液サンプルをpJOrf3-29に対するIgM及びIgG抗体について検査した。図1に示すように、IgM抗体は感染後(P I)51日目に検出され、ALT値がピークに達する72日目まで評価し続けた。pJOrf3-29に対するIgG抗体は最初P I56日目に検出され、107日目まで陽性のままであった(表50)。

【0271】

HEV US ORF3を表すがCKS融合パートナーを欠く第2構築物p1orf3-12も上記と同一のELISAフォーマットで評価した。p1orf3-12に対するIgG抗体を同様に実験的に感染させたマカクからの一連の血液で評価した。p1orf3-12に対するIgG抗体がP I58日目に検出され、107日目まで陽性のままであった(表50)。

【0272】

【表49】

表50

サンプル	pJ0orf3-29		p1orf3-12	
	平均OD	S/N	平均OD	S/N
SpD			0.01	
接種前血液	0.02		0.01	
接種後出血				
DPI				
44	0.02	0.96	0.02	1.07
51	0.05	2.35	0.03	2.25
56	0.24	10.35	0.05	3.43
58	0.44	19	0.16	11.57
63	1.14	49.57	0.32	22.82
65		NT	0.53	37.54
70		NT	1.19	85.04
72	2.22	96.52	0.92	65.71
98	0.89	38.87	0.39	27.86
107	0.49	21.43	0.27	19.36

DPI：接種後の日数

NT：検査せず

【0273】

ブタH E V及びU S - 2単離物間の相同%が高いために、p J O o r f 3 - 2
9 E L I S Aも米国のブタから単離した血清中の免疫反応性I g G及びI g M
の両方の罹患率を調べるために使用した(表5 1)。このアッセイは上記したよ
うに実施したが、ただし抗 - ヒトコンジュゲートの代わりにH P R O - コンジュ
ゲートした標識抗 - ブタ免疫グロブリン (I g GまたはI g M) を使用した。

【 0 2 7 4 】

【表5 0】

表51

米国ブタにおけるHEV orf3に対する抗体保有率 (pJOrf3-29)					
ブタソース	I g G反応性の 数/総数 (%)	ブロッキング またはブロット で確認された I g G数 (%)	I g Mのみ 反応性の数/総数 (%)	プロットで 確認された I g Mのみの数 (%)	確認のみされた 総接触
ニュージャージー	9/14 (64)	9 (100)	0/14		64%
テキサス	25/ (50)	20 (80)	0/50		40%
アイオワ	7/64 (11)	1 (14)	0/64		2%
オレゴン	7/36 (19)	5 (71)	1/36 (3)	1/1 (100)	14%
総数	48/164 (29)	35 (73)	1/164 (0.6)	1/1 (100)	36/164 (22%)

注：4匹のブタ（テキサス群のすべて）はI g G及びI g Mを有していた。

【0275】

反応性試料を確認するために、ブロッキングアッセイを実施した。簡単に説明すると、予備希釈した1：16試料の10μlアリコート反応トレーの2つの

ウェルに添加した。1つのウェルは標準アッセイのために使用し、もう1つのウェルはブロッキングアッセイのために使用した。標準アッセイのためのE L I S Aは、p J O o r f 3 - 2 9抗原コートしたビーズを添加する前に室温で30分間ブレインキュベートするステップを含めた以外は上記のように実施した。ブロッキングアッセイのために、p J O o r f 3 - 2 9をS p D (ブロッキング試薬) に対して固相上よりも10倍モル過剰添加した。1反応あたり200 μ lのブロッキング試薬を添加し、p J O o r f 3 - 2 9抗原コートしたビーズを添加する前に室温で30分間ブレインキュベートした。アッセイの残りは、抗 - ヒトコンジュゲートの代わりにH P R O - コンジュゲートした抗 - ブタコンジュゲート (I g G) を使用した以外はブタアッセイに関して上記したように実施した。

【0276】

ブロッキング率 (%) は、式：

$$\left[\left(A_{492nm}^{\text{標準アッセイ}} - A_{492nm}^{\text{ブロッキングアッセイ}} \right) / A_{492nm}^{\text{標準アッセイ}} \right] \times 100$$

を用いて求めた。50%以上のブロッキング率を示した試料がH E V p J O o r f 3 - 2 9に対するI g G抗体に対して反応性であると見做された。代表的なI g G陽性及びI g G陰性ブタサンプル及びブロッキング結果を表52に示す。

【0277】

【表51】

表52：10倍モル過剰で pJOorf3-29 及び PL-12 を用いる
ブロッキングアッセイ

標準アッセイ			10倍モル過剰で pJOorf3-29 を 用いるブロッキングアッセイ			ブロッキング 結果
サンプル	OD	平均 OD	OD	平均 OD	ブロッキング %	
NC	0.02 0.02 1.09	0.02	0.02 0.03 0.56	0.02		
PC	1.01	1.05	0.48	0.52	50.4%	+
オレゴンブタパネルポジティブ						
1	NJ5	0.65	0.15		76.5%	+
2	NJ12	1.78	0.46		74.0%	+
3	NJ21	0.48	0.16		66.7%	+
4	NJ23	0.52	0.09		81.9%	+
5	T5	2	0.81		59.5%	+
6	T9	0.52	0.18		64.3%	+
7	T32	2	0.9		54.9%	+
8	T33	0.3	0.13		57.8%	+
9	T48	0.53	0.14		73.7%	+
10	T49	0.33	0.09		73.3%	+
オレゴンブタパネルネガティブ						
11	T43	0.08	0.07		13.3%	-
12	T46	0.12	0.08		29.1%	-
13	I-23	0.12	0.08		32.2%	-
14	I-24	0.07	0.06		13.2%	-
15	I-27	0.1	0.08		12.6%	-
16	I-28	0.15	0.12		20.4%	-
17	I-33	0.15	0.12		19.9%	-
18	I-39	0.23	0.14		37.4%	-
19	I-61	0.19	0.14		25.9%	-
20	O-4	0.15	0.12		22.7%	-

【0278】

ブロッキングアッセイに加えて、別の組のブタ試料に対してウェスタンブロットを実施した。簡単に説明すると、HEV pJOorf3-29 (50 µg) 及び“CKSのみ”タンパク質 (50 µg) をSDS-PAGEで分画化し、分画化したタンパク質をニトロセルロースに移した。ニトロセルロースの3mmストリップを切断し、環状ローターにおいて室温で10%大腸菌ライゼートを含むタンパク質ベース緩衝液において1:100希釈で一次抗体と一晚インキュベートした。翌日、ストリップを0.3% ツイーン/TBST (TBST) で3回洗浄後TBSTで0.5 µg/mlに希釈したHPRO-コンジュゲートした抗-ブタ免疫IgGコンジュゲートを添加した。ストリップを室温で回転させながら4時間インキュベートした。次いで、プロットをTBSTで3回洗浄した後TBSTで2回洗浄した。プロットを基質として4-クロロ-1-ナフトールを用いて展開させた。水を添加して反応を停止し、バンド強度を記録した。試料がp

J O o r f 3 - 2 9 (約 4 0 k D) に関して正確な分子量のバンドを示し、“ C K S のみ ” バンド (約 2 9 k D) を含む領域では反応性がない場合には H E V に対して特異的反応性を有すると判定された。 p J O o r f 3 - 2 9 ウェスタンブロットに対して実施した 2 0 個のブタ血清の結果を表 5 3 に示す。ブタ血清は、“ C K S のみ ” バンドと非特異的反応性を示さなかった。

【 0 2 7 9 】

【表 5 2】

表 5 3

ブタ I D 番号	バンド強度	
	pJ0rf3-29	C K S のみ
N J 4	+	—
N J 7	+	—
N J 1 4	+++	—
N J 1 8	+	—
N J 2 5	++++	—
T 6	++++	—
T 1 0	++++	—
T 1 4	—	—
T 1 5	+	—
T 1 8	++	—
T 2 8	+++	—
T 2 9	—	—
T 3 0	+	—
T 3 4	—	—
T 3 6	++++	—
T 3 7	—	—
T 4 3	—	—
T 4 4	++++	—
T 4 5	++++	—
T 4 6	—	—

【 0 2 8 0 】

これらのデータから、 H E V U S 組み換えタンパク質が H E N 接触を診断する際に有用であることが示唆される。

【 0 2 8 1 】

実施例 1 1 : コンセンサスプライマー

HEV ORF 1、ORF 2及びORF 3に対するコンセンサスオリゴヌクレオチドプライマーをアジア、メキシコ及び米国由来の単離物の完全長配列間の保存領域に基づいて設計した(図9)。ORF 1プライマーはビルマ単離物(GenBank受託番号M73218)のヌクレオチド56~79及び473~451のメチルトランスフェラーゼ領域内に位置しており、418ヌクレオチド長の産物を増幅させる。ORF 1プライマーには、

HEVConsORF1-s1;CTGGCATYACTACTGCYATTGAGC(配列番号147)及び

HEVConsORF1-a1;CCATCRARRCAGTAAGTGCGGTTC(配列番号148)

が含まれる。

【0282】

ビルマ単離物の6298~6321及び6494~6470位置のORFプライマーは197ヌクレオチド長の産物を生ずる。ORF 2プライマーには、

HEVConsORF2-s1;GACAGAATTTRATTTTCGTCTGGCTGG(配列番号150)及び

HEVConsORF2-a1;CTTGTTTCRTGYTGGTTTRTCATAATC(配列番号126)

が含まれる。

【0283】

第2ラウンドの増幅の場合、ORF 1及びORF 2に対してそれぞれ287及び145ヌクレオチド長の産物を生じさせるために内部プライマーを使用した。

ORF 1プライマーには、

HEVConsORF1-s2;CTGCCYTKGCGAATGCTGTGG(配列番号177)及び

HEVConsORF1-a2;GGCAGWRTACCARCGCTGAACAATC(配列番号178)

が含まれる。

【0284】

ORF 2 プライマーには、
 HEV Cons ORF 2 - s 2 ; G T Y G T C T C R G C C A A T G G C G A G
 C (配列番号 1 5 2) 及び
 HEV Cons ORF 2 - s 2 ; G T T C R T G Y T G G T T R T C A T A A T
 C C T G (配列番号 1 2 8)
 が含まれる。

【0285】

PCR 反応物は製造業者 (P e r k i n - E l m e r) の指示に従って 2 m M の M g C l ₂ 及び 0 . 5 μ M の各オリゴヌクレオチドプライマーを含め、実施例 5 に記載のタッチダウン PCR を用いて増幅した。増幅した産物を 1 . 5 % アガロースゲルを用いて分離し、適当なサイズを有する PCR 産物の存在について分析した。このプライマーを、HEV US - 2 を含む血清及び糞中のウイルスの存在を実施例 8 及び図 7 に記載されているように検出するために用いた。更に、前記プライマーはビルマ様株 6 A、7 A、9 A 及び 1 2 A を含めた HEV の多種多様変異体、ギリシャ由来の 2 つの異なる単離物 (以下の実施例 1 3 参照)、イタリア由来のユニークな単離物及び米国由来の 2 つの単離物 (以下の実施例 1 3 参照) と反応することが判明した。更に、前記プライマーを使用して、中国の遼寧省の急性散在性肝炎と臨床診断された患者 (S 1 5) からの単離物を同定した。結果を下表 5 4 に示す。

【0286】

【表 5 3】

表 5 4

サンプル	ORF 1 -PCR1	ORF 1 -PCR 2	ORF 2 - PCR1	ORF 2 -PCR2
6A	neg	pos	pos	Pos
7A	neg	pos	neg	Pos
9A	neg	neg	neg	Pos
12A	pos	pos	neg	Neg
G1	pos	pos	pos	Pos
G2	pos	pos	pos	Pos
Itl	pos	pos	pos	Pos
S15	nd	pos	nd	Pos
US-2	pos	pos	pos	Pos

【0287】

実施例12：一次ヒト胎児腎細胞におけるHEV RNAの検出

1.0×10^6 細胞を含む凍結細胞ペレットを融解し、ダルベッコリン酸塩緩衝食塩液(1.0 ml)に再懸濁した。RNAを細胞ペレット($20 \mu\text{l}$; 2×10^5 細胞)から実施例1に記載のUltraspec Isolation Systemを用いて抽出した。cDNA合成を上記のように抽出した核酸(RNA)に対して実施、ランダムヘキサマーを用いてプライムした。次いで、上記cDNAに対して、実施例11に記載のように $0.5 \mu\text{M}$ の最終濃度でウイルスゲノムのORF-1及びORF-2領域由来の縮重プライマーを用いてPCRを実施した。

【0288】

上記アッセイの性能をモニターするために、一次ヒト腎細胞及びHEV US-2陽性血清を用いるポジティブコントロールを実験計画に含めた。2つのポジティブコントロールを、 2×10^5 個のHEV陰性一次ヒト腎細胞を公知のHEV US-2陽性血清標本($2.5 \mu\text{l}$ 及び $25 \mu\text{l}$)でスパイクすることにより調製した。ポジティブコントロール血清もヒト腎細胞を添加せずに試験した。

【0289】

一次ヒト腎細胞ペレットの19個のロットをORF1及びORF2由来の2つの縮合プライマー組を用いて上記アッセイで試験した。結果を下表55に要約する。試験した細胞ペレットのいずれのロットもポジティブコントロールでみられるような陽性結果を示さなかった。

【0290】

【表54】

表55

細胞系	P C R 結果
1 7 5 7	—
1 8 5 1	—
1 6 9 0	—
1 8 5 3	—
1 9 0 6	—
1 9 3 5	—
1 8 3 8	—
1 9 5 5	—
1 8 9 3	—
1 8 9 5	—
1 6 9 9	—
1 8 7 7	—
1 9 4 2	—
1 8 4 4	—
1 8 4 0	—
1 8 7 5	—
1 9 2 1	—
1 9 4 6	—
1 8 4 6	—
細胞+25 μ l 血清	+
細胞+2.5 μ l 血清	+
25 μ l 血清	+

【0291】

実施例13：別のUSタイプ単離物の同定及び延長

(A) イタリア単離物 (It1と称する) の同定

RNAを血清(25~50 μ l)からQAI Amp Viral RNAキット(Qiagen)を製造業者が記述しているように用いて抽出した。ただし、血清(25~50 μ l)をPBS(100 μ l)で希釈し、最終溶離をRNase非含有水(100 μ l)で実施した。RT反応はランダムプライムした。PCRは実施例5に上記したHEV US-1プライマーを用いた。プライマー(配列番号94及び配列番号96)を用いて増幅後294bp産物が生じた。この産物を実施例3に記載のようにクローン化し、配列決定し、配列番号179に示す。

【0292】

I t 1 単離物を以下のように延長させた。RNAを実施例5に上記されているように血清(25~50 µl)から抽出した。RT反応はランダムプライムした。PCRは実施例11に上記したHEV CONSENSUSプライマーを実施例3に上記したようにタッチダウンPCRにおいて用いた。配列番号147及び148に示すプライマーを用いて、配列番号180に示す配列を有する産物を生成した(反応z2, 418 bp)。配列番号150及び126に示すプライマーを用いて、配列番号181に示す配列を有する産物を生成した(反応z3, 197 bp)。1×PCR緩衝液及び20% Q Solution(Qiagen)の存在下で、配列番号182及び183に示すプライマーを用いて配列番号184に示す配列を有する産物を生成した(反応z4, 234 bp)。ゲノムの3'末端をPCR1では配列番号150及び85に示すプライマー、PCR2では配列番号152及び85に示すプライマーを用いて実施例3に上記した3' RACEにより単離して、配列番号185に示す配列を有する産物を生成した(反応z5, 890 bp)。産物を実施例3に記載のようにクローン化し、配列決定し、コンセンサス配列を作成した。これらの領域を図8に示し、配列番号180、184及び186に示す、前記領域のアミノ酸翻訳を配列番号187、188、189、190及び197に示すアミノ酸配列で表す。

【0293】

(B) 2つのギリシャ単離物(G1及びG2と呼ぶ)の同定

病気が流行している地域へ旅行したことがない2人の急性肝炎患者をビルマ単離物(Psichogiou M.A.ら,「急性非A非B型肝炎患者のコホートにおけるE型肝炎ウイルス(HEV)感染(Hepatitis E virus (HEV) infection in a cohort of patients with acute non-A, non-B hepatitis)」, Journal of Hepatology, 23:668-673(1995))に基づくプライマーを用いて分析した。患者G2のみがPCR陽性であると判明した。RNAを実施例12に上記したように単離し、PCRを実施例11に上記のコンセンサスプライマーを用いて実施した。ORF1及びORF2プライマー組により、両方の患者から予想されるサイズを有する産物が生成した。この産物

を実施例3に上記したようにクローン化し、配列決定した。患者G1からORF1及びORF2コンセンサプライマーを用いて生成した産物をそれぞれ配列番号209及び211に示す。患者G2からORF1及びORF2コンセンサプライマーを用いて生成した産物をそれぞれ配列番号213及び215に示す。G1がPCR陽性であると同定されたことから、コンセンサプライマーがビルマ株特異的プライマーよりも有用であることが立証される。

【0294】

G1及びG2由来の追加配列も、実施例3に上記した配列番号19の作成の場合のようにプライマー（配列番号16、配列番号17及び配列番号18）を用いて得た。ただし、PCRのためにランダムプライマーcDNAを用い、増幅には94 × 20秒、60 × 30秒及び72 × 1分の10サイクル、その後に94 × 20秒、55 × 30秒及び72 × 1分の10サイクル及びその後に94 × 20秒、50 × 30秒（-0.3 / サイクル）及び72 × 1分の30サイクルを含めた。この後に、72 × 7分の延長サイクルを続けた。患者G1から生成した産物を配列番号217に示す。患者G2から生成した産物を配列番号220に示す。

【0295】

米国、中国、ギリシャ、イタリア、メキシコ及びビルマ様単離物のヌクレオチド配列のアラインメントを前記単離物の相互の関係を調べるために実施した。イタリア単離株の相違は、ゲノムのORF1領域由来の産物をビルマ（B1）、メキシコ（M1）及び米国（US-1）由来の原型単離物と比較したときの核酸同一性がそれぞれ77.6%、78.4%及び84.6%であることにより裏付けられる（表36）。イタリア単離株の相違は、ゲノムのORF2領域由来の産物をビルマ、メキシコ及び米国由来の原型単離物と比較したときの核酸同一性がそれぞれ83.3%、79.7%及び87.8%であることによっても裏付けられる（表37）。ビルマ、メキシコ及び米国由来の原型単離物間のヌクレオチド同一性は両領域の範囲で75.5～82.4%である。これらの同一領域の範囲で、ビルマ様グループを含む単離物は91.2%以上の非常に高い同一性を有する。

【0296】

ORF 1 及び ORF 2 増幅配列の比較から、2 人のギリシャ人患者からの単離物は相互に全く異なっており、ORF 1 及び ORF 2 領域の範囲でそれぞれ 84.4% 及び 87.2% のヌクレオチド配列同一性を示すことが分かる。ヌクレオチドレベルで、ギリシャ、イタリア及び米国単離物間の同一性は ORF 1 産物に対して 81.9 ~ 86.8% (表 36)、ORF 2 産物に対して 82.4 ~ 87.8% (表 37) である。これらの値は、ORF 1 及び ORF 2 の両方に対するビルマ様単離物間のヌクレオチド同一性の最低%が 91.2% 以上であるのに比べて低い。米国、イタリアまたはギリシャ単離物とビルマまたはメキシコ単離物間の ORF 1 断片由来のアミノ酸同一性を比較すると、87.8 ~ 93.5% である (表 36)。これらの値は、ビルマ単離物とメキシコ単離物間の相違 (93.5 ~ 95.1%) (表 36) と同等またはそれ以下であり、非流行地域由来の単離物は流行地域を起源とする単離物とは区別されることが分かる。

【0297】

分析したウイルス配列間の相対的進化距離は、1 対距離から作成した根無し系統樹の精査から明白である。ここでは、枝の長さが単離物間の相対的遺伝的關係に比例している。ORF 1 (図 10) または ORF 2 (図 11) 配列のアラインメントに基づく系統樹は全トポロジーにおいて全く同一である。ビルマ様単離物及びメキシコ単離物は樹の 1 端の主要枝を表す。米国人単離物はメキシコ単離物及びビルマ単離物から離れた別のグループを形成する。ORF 2 由来のブタ H E V 様配列は米国人単離物と密接に関連している。3 つのヨーロッパ単離物は 3 つの別の異なる枝を形成し、イタリア単離物が米国単離物に最も密接に関連している。

【0298】

実施例 14：オーストリア及びアルゼンチン由来の別の U S タイプ単離物の同定

H E V が流行していると思われる地域に旅行したことがない 3 人の急性肝炎患者の血清から R N A を実施例 12 に上記したように単離し、P C R を実施例 11 に上記したコンセンサスプライマーを用いて実施した。1 人はオーストリア患者

Au1 (Wormら, 「オーストリアにおける散发性E型肝炎(Sporadic hepatitis E in Austria)」, New England Journal of Medicine, 339:1554-1555(1998))であり、他の2人はアルゼンチン患者であった。ORF1及びORF2プライマー組を用いて、すべての患者から予想されるサイズの産物が生成された。この産物を実施例3に上記したようにクローン化し、配列決定した。患者Au1からORF1及びORF2コンセンサスプライマーを用いて生成した産物をそれぞれ配列番号243及び245に示す。患者Ar1からORF1及びORF2コンセンサスプライマーを用いて生成した産物をそれぞれ配列番号247及び249に示す。患者Ar2からORF1及びORF2コンセンサスプライマーを用いて生成した産物をそれぞれ配列番号251及び253に示す。Au1、Ar1及びAr2に対してa1及びs1プライマーを用いるORF1 PCRの第1ラウンド及びa2及びs2プライマーを用いるネステッドORF1 PCRの第2ラウンド後PCR産物が得られた。Au1及びAr2に対してa1及びs1プライマーを用いるORF2 PCRの第1ラウンド及びa2及びs2プライマーを用いるネステッドORF2 PCRの第2ラウンド後PCR産物が得られた。Ar1からの産物は、a2及びs2プライマーを用いるネステッドORF2 PCRの第2ラウンド後にのみ検出された。

【0299】

米国、中国、ギリシャ、イタリア、オーストリア、アルゼンチン、メキシコ及びビルマ様単離物のヌクレオチド配列のアラインメントを実施例6に記載されているように実施して、前記単離物の相互の関係を調べた。オーストリア単離物Au1の相違は、ゲノムのORF1領域由来の産物をビルマ(B1)、メキシコ(M1)及び米国(US-1)由来の原型単離物と比較したときの核酸同一性がそれぞれ77.1%、78.2%及び87.9%であることにより裏付けられる(表56)。オーストリア単離物の相違は、ゲノムのORF2領域由来の産物をビルマ(B1)、メキシコ(M1)及び米国(US-1)由来の原型単離物と比較したときの核酸同一性がそれぞれ85.1%、79.1%及び83.1%であることによっても裏付けられる(表57)。アルゼンチン単離物Ar2の相違は、

ゲノムのORF 1領域由来の産物をビルマ (B 1)、メキシコ (M 1) 及び米国 (U S - 1) 由来の原型単離物と比較したときの核酸同一性がそれぞれ76.0%、76.0%及び84.9%であることにより裏付けられる (表56)。Ar 2単離物の相違は、ゲノムのORF 2領域由来の産物をビルマ (B 1)、メキシコ (M 1) 及び米国 (U S - 1) 由来の原型単離物と比較したときの核酸同一性がそれぞれ85.8%、82.4%及び85.8%であることにより裏付けられる (表57)。アルゼンチン単離物Ar 1の相違は、ゲノムのORF 1領域由来の産物をビルマ (B 1)、メキシコ (M 1) 及び米国 (U S - 1) 由来の原型単離物と比較したときの核酸同一性がそれぞれ76.6%、77.6%及び85.7%であることにより裏付けられる (表56)。ビルマ (B 1)、メキシコ (M 1) 及び米国 (U S - 1) 由来の原型単離物間のヌクレオチド同一性は2つの領域の範囲で75.5~82.4%である。同一領域の範囲で、ビルマ様グループを含む単離物は91.2%以上という非常に高い同一性を有する。ネステッドORF 2 PCR産物のみがアルゼンチン単離物Ar 1から得られたが、Ar 2単離物の相違はゲノムのORF 2領域由来の小産物をビルマ (B 1)、メキシコ (M 1) 及び米国 (U S - 1) 由来の原型単離物と比較したときの核酸同一性が80.6%であることにより裏付けられる (表57)。

【0300】

ヌクレオチドレベルで、(U S - 1及びU S - 2間の同一性を除き) オーストリア、アルゼンチン、ギリシャ、イタリア及び米国単離物間の同一性はORF 1産物に対して80.6~89.8%である (表56)。ヌクレオチドレベルで、(U S - 1、U S - 2、Ar - 1及びAr - 2間の同一性を除き) オーストリア、アルゼンチン、ギリシャ、イタリア及び米国単離物間の同一性はORF 2産物に対して80.6~89.2%である (表57)。これらの値は、ORF 1及びORF 2に対するビルマ様単離物間のヌクレオチド同一性の最低%が91.2%あるのに比べて低い。

【0301】

【表55】

表 56

371塩基(123アミノ酸)ORF1断片の範囲のHEV単離株間のヌクレオチド同一性及び推定アミノ酸同一性

ヌクレオチド同一性																				
Ar1	88.4	89.8	84.4	81.9	85.4	85.7	85.2	82.5	76.6	76.6	79.0	78.2	79.2	77.1	78.4	76.8	78.7	77.6		
98.4	Ar2	87.9	80.6	81.1	84.4	84.9	85.4	83.3	76.0	76.6	74.9	75.7	75.7	74.1	77.4	76.8	75.7	76.0		
100	98.4	An1	85.2	81.1	86.0	87.9	87.1	84.6	77.1	77.6	76.8	76.6	77.6	76.6	77.4	78.2	77.1	78.2		
99.1	97.6	99.1	G1	84.4	84.1	81.9	82.5	83.0	78.4	77.9	77.6	78.2	77.9	76.6	77.6	78.2	77.4	76.6		
98.4	96.7	98.4	99.2	G2	81.7	83.8	83.0	81.9	78.2	77.6	78.2	77.9	78.4	77.1	77.9	77.6	78.7	76.8		
97.6	95.9	97.6	96.7	It1	84.6	86.8	84.9	77.6	77.6	77.6	77.1	77.4	77.4	76.3	77.6	77.4	77.6	78.4		
99.2	97.6	99.2	98.4	97.6	96.7	US-1	91.9	90.8	75.5	74.9	75.2	75.2	75.7	75.2	76.6	75.5	76.0	76.6		
100	98.4	100	99.2	98.4	97.6	99.2	US-2	89.9	75.2	75.4	75.2	75.4	76.0	74.9	75.7	75.7	76.3	77.6		
97.6	95.9	97.6	96.7	95.9	95.1	96.7	97.6	S1	76.6	76.6	74.7	74.9	74.4	73.6	75.2	74.9	75.2	75.7		
91.1	90.2	91.1	90.2	90.2	92.7	91.9	91.1	88.6	B1	98.7	94.6	94.6	94.9	94.3	94.3	96.0	94.6	79.0		
91.1	90.2	91.1	90.2	90.2	92.7	90.2	91.1	89.6	98.4	B2	93.8	93.8	94.1	93.5	93.5	95.1	93.8	78.4		
89.4	88.6	89.4	88.6	88.6	91.1	90.2	89.4	87.0	98.4	96.7	C1	97.8	98.1	96.8	92.7	91.6	97.3	79.8		
90.2	89.4	90.2	89.4	89.4	91.9	91.1	90.2	87.8	99.2	97.6	99.2	C2	98.7	97.6	93.8	91.6	98.4	79.5		
90.2	89.4	90.2	89.4	89.4	91.9	91.1	90.2	87.8	99.2	97.6	99.2	100	C3	97.3	93.5	91.9	98.1	79.5		
89.4	88.6	89.4	88.6	88.6	91.9	90.2	89.4	87.0	98.4	96.7	99.2	99.2	99.2	C4	93.0	91.9	97.0	78.7		
90.2	89.4	90.2	89.4	89.4	91.9	91.1	90.2	87.8	99.2	97.6	97.6	98.4	98.4	97.6	I1	91.4	93.3	79.2		
88.6	87.8	88.6	87.8	87.8	90.2	89.4	88.6	86.2	97.6	95.9	95.9	96.7	96.7	96.7	I2	91.6	78.4	78.4		
90.2	89.4	90.2	89.4	89.4	91.9	91.1	90.2	87.8	99.2	97.6	99.2	100	100	99.2	98.4	96.7	P1	78.7		
97.7	91.1	92.7	91.9	91.9	94.3	93.5	92.7	90.2	95.9	95.1	94.3	95.1	95.1	94.3	93.5	95.1	95.1	M1		

アミノ酸同一性

【0302】

【表56】

表 57

148 塩基 (49 アミノ酸) * ORF2断片の範囲の HEV 単離株間のヌクレオチド及び推定アミノ酸同一性

		ヌクレオチド同一性																アミノ酸同一性													
		Ar1	91.8	87.8	81.6	82.7	83.7	80.6	82.7	87.8	80.6	80.6	80.6	80.6	80.6	80.6	80.6	80.6	80.6	80.6	80.6	80.6	80.6	80.6	80.6	80.6	80.6	80.6	80.6	80.6	80.6
100	Ar2	100	88.5	83.8	86.5	87.2	85.8	85.1	90.5	85.8	85.1	83.8	85.1	85.1	85.1	85.1	85.1	85.1	85.1	85.1	85.1	85.1	85.1	85.1	85.1	85.1	85.1	85.1	85.1	85.1	85.1
100	An1	100	83.1	88.5	89.2	83.1	85.8	87.8	85.1	85.1	84.5	82.4	82.4	83.1	83.8	83.8	83.1	83.8	83.1	83.8	83.1	83.1	82.4	83.1	82.4	83.1	82.4	83.1	82.4	83.1	82.4
100	G1	100	100	100	87.2	87.8	84.5	85.1	85.1	85.1	84.5	82.4	82.4	83.1	83.8	83.8	83.1	83.8	83.1	83.8	83.1	83.1	82.4	83.1	82.4	83.1	82.4	83.1	82.4	83.1	82.4
100	G2	100	100	100	100	83.1	82.4	85.1	87.8	85.1	84.5	82.4	82.4	83.8	83.8	83.8	83.1	83.8	83.1	83.8	83.1	83.1	82.4	83.1	82.4	83.1	82.4	83.1	82.4	83.1	82.4
100	It1	100	100	100	100	100	87.8	85.8	85.8	85.8	83.8	83.1	82.4	83.1	83.1	83.1	83.1	83.1	83.1	83.1	83.1	83.1	83.1	83.1	83.1	83.1	83.1	83.1	83.1	83.1	83.1
96.9	US-1	98.0	98.0	98.0	98.0	98.0	98.0	98.0	93.9	90.5	79.0	78.4	76.4	77.0	77.0	77.0	77.0	77.0	77.0	77.0	77.0	77.0	77.0	77.0	77.0	77.0	77.0	77.0	77.0	77.0	77.0
96.9	US-2	98.0	98.0	98.0	98.0	98.0	98.0	100	100	91.2	82.4	80.4	79.7	80.4	80.4	80.4	80.4	80.4	80.4	80.4	80.4	80.4	80.4	80.4	80.4	80.4	80.4	80.4	80.4	80.4	80.4
96.9	S1	98.0	98.0	98.0	98.0	98.0	98.0	100	100	91.2	82.4	80.4	79.7	80.4	80.4	80.4	80.4	80.4	80.4	80.4	80.4	80.4	80.4	80.4	80.4	80.4	80.4	80.4	80.4	80.4	80.4
96.9	B1	98.0	98.0	98.0	98.0	98.0	98.0	98.0	93.9	93.9	98.0	98.0	98.0	98.0	98.0	98.0	98.0	98.0	98.0	98.0	98.0	98.0	98.0	98.0	98.0	98.0	98.0	98.0	98.0	98.0	98.0
96.9	B2	98.0	98.0	98.0	98.0	98.0	98.0	98.0	93.9	93.9	98.0	98.0	98.0	98.0	98.0	98.0	98.0	98.0	98.0	98.0	98.0	98.0	98.0	98.0	98.0	98.0	98.0	98.0	98.0	98.0	98.0
96.9	C1	98.0	98.0	98.0	98.0	98.0	98.0	98.0	93.9	93.9	98.0	98.0	98.0	98.0	98.0	98.0	98.0	98.0	98.0	98.0	98.0	98.0	98.0	98.0	98.0	98.0	98.0	98.0	98.0	98.0	98.0
96.9	C2	98.0	98.0	98.0	98.0	98.0	98.0	98.0	93.9	93.9	98.0	98.0	98.0	98.0	98.0	98.0	98.0	98.0	98.0	98.0	98.0	98.0	98.0	98.0	98.0	98.0	98.0	98.0	98.0	98.0	98.0
96.9	C3	98.0	98.0	98.0	98.0	98.0	98.0	98.0	93.9	93.9	98.0	98.0	98.0	98.0	98.0	98.0	98.0	98.0	98.0	98.0	98.0	98.0	98.0	98.0	98.0	98.0	98.0	98.0	98.0	98.0	98.0
96.9	C4	98.0	98.0	98.0	98.0	98.0	98.0	98.0	93.9	93.9	98.0	98.0	98.0	98.0	98.0	98.0	98.0	98.0	98.0	98.0	98.0	98.0	98.0	98.0	98.0	98.0	98.0	98.0	98.0	98.0	98.0
96.9	H1	98.0	98.0	98.0	98.0	98.0	98.0	98.0	93.9	93.9	98.0	98.0	98.0	98.0	98.0	98.0	98.0	98.0	98.0	98.0	98.0	98.0	98.0	98.0	98.0	98.0	98.0	98.0	98.0	98.0	98.0
93.8	I2	95.9	95.9	95.9	95.9	95.9	95.9	93.9	93.9	93.9	98.0	95.9	98.0	98.0	98.0	98.0	98.0	98.0	98.0	98.0	98.0	98.0	98.0	98.0	98.0	98.0	98.0	98.0	98.0	98.0	98.0
96.9	P1	98.0	98.0	98.0	98.0	98.0	98.0	98.0	95.9	95.9	98.0	98.0	98.0	98.0	98.0	98.0	98.0	98.0	98.0	98.0	98.0	98.0	98.0	98.0	98.0	98.0	98.0	98.0	98.0	98.0	98.0
96.9	M1	98.0	98.0	98.0	98.0	98.0	98.0	98.0	95.9	95.9	98.0	98.0	98.0	98.0	98.0	98.0	98.0	98.0	98.0	98.0	98.0	98.0	98.0	98.0	98.0	98.0	98.0	98.0	98.0	98.0	98.0

* Ar1 に対しては 98 塩基 (32 アミノ酸) 断片の範囲

【0303】

ORF 1 及び ORF 2 増幅配列の比較から、2 人のアルゼンチン患者からの単離物は相互に全く異なり、ORF 1 及び ORF 2 領域の範囲でそれぞれ 88.4% 及び 91.4% のヌクレオチド配列同一性を示すことが分かる。ORF 1 に対する値は、ORF 1 に対してビルマ様単離物間のヌクレオチド同一性の最低%が 91.4% であるのに比べて低い。しかしながら、ORF 2 の場合、91.8% というアルゼンチン由来の 2 つの単離物間のヌクレオチド同一性は多分断片の長

さがより短いためにビルマ様単離物とORF 2間の同一性で見られる範囲内である。

【0304】

系統発生的分析を実施例7に記載されているように実施した。分析したウイルス配列間の相対的進化距離は1対距離から作成した根無し系統樹を精査することにより明白である。ここで、枝の長さは単離物間の相対的遺伝的關係に比例する。ORF 1由来の371ヌクレオチド(図14)、Ar 1を除くORF 2由来の148ヌクレオチド(図15)またはAr 1を含めたORF 2由来の98ヌクレオチド(図16)のアラインメントに基づく系統樹は全トポロジーにおいて非常に類似している。ビルマ様単離物及びメキシコ単離物は樹の1端で主要枝を表す。米国人単離物はメキシコ及びビルマ単離物から離れた別のグループを形成する。ブタHEV様配列は米国人単離物と密接に関連している。4つのヨーロッパ単離物及び2つのアルゼンチン単離物もメキシコ単離物及びビルマ単離物から離れた枝を形成する。米国、ギリシャ、イタリア、オーストリア及びアルゼンチン単離物で表される米国タイプの単離物とビルマ様及びメキシコ単離物間の主要枝は、すべての樹において75.7%以上のブーツストラップ値により裏付けられる。

【0305】

実施例15：新規な縮合プライマー

HEV ORF 1及びORF 2に対するコンセンサスオリゴヌクレオチドプライマーから誘導した縮合プライマーを、実施例11に記載のアジア、メキシコ及び米国由来の単離物とギリシャ及びイタリア由来の単離物の完全長配列間の保存領域に基づいて設計した。ORF 1プライマーはビルマ単離物(GenBank受託番号M73218)のヌクレオチド473~451のメチルトランスフェラーゼ領域内に位置し、実施例11に記載のHEVConsORF 1-s1(配列番号147)と組合せて使用したときに417ヌクレオチド長の産物を増幅する。新規なORF 1プライマーの組合せには、

HEVConsORF 1-s1; CTGGCATYACTACTGCYATTGAGC(配列番号147)及び

HEVConsORF1N-a1;CCRTCRARRRCARTAGGTGCG
GTC(配列番号255)

が含まれる。

【0306】

ビルマ単離株の6494~6470の位置の新規なORF2プライマーは実施例11に記載のHEVConsORF2-s1(配列番号150)と組合せて使用したときに197ヌクレオチド長の産物を生成する。新規なORF2プライマーには、

HEVConsORF2-s1;GACAGAATTTRATTTTCGTCTGGC
TGG(配列番号150)及び

HEVConsORF2N-a1;CYTGYTCRTGYTGGTTTCA
TAATC(配列番号256)

が含まれる。

【0307】

第2ラウンドの増幅の場合には、実施例11に記載のようにORF1及びORF2に対してそれぞれ287及び145ヌクレオチド長の産物を生成するために内部プライマーを使用することができる。ORF1プライマーの新規な組合せには、

HEVConsORF1N-s2;CYGCCYTKGCGAATGCTGTG
G(配列番号257)及び

HEVConsORF1-a2;GGCAGWRTACCARCGCTGAAC
ATC(配列番号178)

が含まれる。

【0308】

ORF2プライマーには、

HEVConsORF2-s2;GTYGTCCTCRGCCAATGGCGAG
C(配列番号152)及び

HEVConsORF2N-a2;GYTCRTGYTGRTTTRTCATAA
TCCTG(配列番号258)

が含まれる。

【0309】

PCR反応物は、製造業者(Perkin-Elmer)の指示に従って2 mMのMgCl₂及び0.5 μMの各プライマーを含み、実施例5に記載のタッチダウンPCRを用いて増幅させた。増幅産物を1.5%アガロースゲル上で分離し、臭化エチジウムで染色し、適当なサイズのPCR産物について分析した。上記したようにHEVを含有する血清中のウイルスの存在を検出するためにプライマーを使用し、実施例11で使用した従来のプライマー組に比して感度の著しい上昇を示した。これらの新規なプライマーの組合せは、2つのアルゼンチン由来の新規単離物Ar1及びAr2、オーストリア由来の新規単離物Au1(上記実施例14参照)、ギリシャ由来の単離物G1及びエジプト由来の単離物Eg46を含めた多種多様のHEV変異体に対してより感受性であることが判明した。結果を下表58に示すが、表中、NTは試験しなかったサンプルを指し、“-”は臭化エチジウム染色により検出可能な産物バンドがないことを指し、“±”は臭化エチジウム染色により検出可能な弱い産物バンドを指し、“2+”、“3+”及び“4+”は臭化エチジウム染色により検出された産物の増加量を指す。

【0310】

【表57】

表58

	ORF1				ORF2			
	PCR1		PCR2		PCR1		PCR2	
	旧組	新組	旧組	新組	旧組	新組	旧組	新組
サンプル								
Ar1	—	2+	2+	4+	2+	4+	3+	4+
Ar2	—	2+	3+	4+	±	±	—	3+
Au1	—	2+	3+	4+	—	3+	3+	4+
Eg46	NT	NT	NT	NT	—	3+	3+	4+
G1	—	—	2+	+	3+	3+	3+	4+

【0311】

均等

本発明は、趣旨またはその必須要件を逸脱しない限り他の特定形態も包含し得る。よって、上記した実施態様は、本明細書に記載の本発明を限定するよりはむ

しるすべての点で例示と見做される。従って、本発明の範囲は以上の記載ではなく請求の範囲により規定されており、請求の範囲の均等の意味及び範囲に入るすべての変更が本発明に包まれると意図される。

【図面の簡単な説明】

【図 1】

ORF 1、ORF 2 及び ORF 3 領域の相対位置を示す HEV ゲノムの概略図である。

【図 2】

入院 1 日目から 37 日目までの患者 USP - 1 中の血清アスパルテートアミノトランスフェラーゼ及び血清総ビリルビンのレベルを示すグラフである。

【図 3】

本研究中に単離したクローンの相対位置を示す HEV US - 1 ゲノムの概略図である。

【図 4】

本研究中に単離したクローンの相対位置を示す HEV US - 2 ゲノムの概略図である。

【図 5】

完全長 HEV US - 1、HEV US - 2 及び 10 個の他の HEV 単離物由来のヌクレオチド配列の関係を示す根なし系統樹を示す。

【図 6】

ORF 2 / 3 領域由来のヌクレオチド配列（すなわち、配列番号 89 のヌクレオチド残基 5094 ~ 7114 に相当する配列）の関係を示す根なし系統樹を示す。

【図 7】

患者 USP - 2 から採取した血清を接種する前またはその後のマカク中のアラニンアミノトランスフェラーゼ、血清アスパルテートトランスフェラーゼ及び γ -グルタミルトランスフェラーゼのレベルを示すグラフである。

【図 8】

本研究中に単離したクローンの相対位置を示す It 1 ゲノム概略図である。

【図9A】

ORF 1、ORF 2 - 3 及びORF 2 に対するHEVコンセンサスプライマーの設計を示すビルマ(B 1)、メキシコ(M 1)、中国(C 1)、パキスタン(P 1) 及びUS - 1のアラインメントを示す。

【図9B】

ORF 1、ORF 2 - 3 及びORF 2 に対するHEVコンセンサスプライマーの設計を示すビルマ(B 1)、メキシコ(M 1)、中国(C 1)、パキスタン(P 1) 及びUS - 1のアラインメントを示す。

【図9C】

ORF 1、ORF 2 - 3 及びORF 2 に対するHEVコンセンサスプライマーの設計を示すビルマ(B 1)、メキシコ(M 1)、中国(C 1)、パキスタン(P 1) 及びUS - 1のアラインメントを示す。

【図9D】

ORF 1、ORF 2 - 3 及びORF 2 に対するHEVコンセンサスプライマーの設計を示すビルマ(B 1)、メキシコ(M 1)、中国(C 1)、パキスタン(P 1) 及びUS - 1のアラインメントを示す。

【図9E】

ORF 1、ORF 2 - 3 及びORF 2 に対するHEVコンセンサスプライマーの設計を示すビルマ(B 1)、メキシコ(M 1)、中国(C 1)、パキスタン(P 1) 及びUS - 1のアラインメントを示す。

【図9F】

ORF 1、ORF 2 - 3 及びORF 2 に対するHEVコンセンサスプライマーの設計を示すビルマ(B 1)、メキシコ(M 1)、中国(C 1)、パキスタン(P 1) 及びUS - 1のアラインメントを示す。

【図9G】

ORF 1、ORF 2 - 3 及びORF 2 に対するHEVコンセンサスプライマーの設計を示すビルマ(B 1)、メキシコ(M 1)、中国(C 1)、パキスタン(P 1) 及びUS - 1のアラインメントを示す。

【図10】

371ヌクレオチド長で配列番号89の残基26～396に相当するORF1
ヌクレオチド配列の関係を示す根なし系統樹を示す。

【図11】

148ヌクレオチド長で配列番号89の残基6307～6454に相当するORF2
ヌクレオチド配列の関係を示す根なし系統樹を示す。

【図12】

好ましいHEV-US組み換えタンパク質構築物の概略図である。

【図13】

患者USP-2から採取した血清を接種する前及びその後のマカク中のアラニン
アミノトランスフェラーゼ、IgG及びIgMのレベルを示すグラフである。

【図14】

371ヌクレオチド長で配列番号89の残基26～396に相当するORF1
ヌクレオチド配列の関係を示す根なし系統樹を示す。

【図15】

148ヌクレオチド長で配列番号89の残基6307～6454に相当するORF2
ヌクレオチド配列の関係を示す根なし系統樹を示す。

【図16】

98ヌクレオチド長で配列番号89の残基6354～6451に相当するORF2
ヌクレオチド配列の関係を示す根なし系統樹を示す。

【配列表】

SEQUENCE LISTING

<110> Abbott Laboratories
 Schlauder, George G.
 Erker, James C.
 Desai, Suresh M.
 Dawson, George J.
 Mushahwar, I. K.

<120> METHODS AND COMPOSITIONS FOR DETECTING
 HEPATITIS E VIRUS

<130> 6232.US.P1

<140> US 09/468,147
 <141> 1999-12-21

<150> US 09/173,141
 <151> 1998-10-15

<150> US 60/061,199
 <151> 1997-10-15

<160> 258

<170> FastSEQ for Windows Version 3.0

<210> 1
 <211> 18
 <212> DNA
 <213> Artificial Sequence

<220>
 <223> Primer C375M

<400> 1
 ctgaacatcc cggccgac 18

<210> 2
 <211> 19
 <212> DNA
 <213> Artificial Sequence

<220>
 <223> Primer A1-350M

<400> 2
 agaaagcagc gatggagga 19

<210> 3
 <211> 21
 <212> DNA
 <213> Artificial Sequence

<220>
 <223> Primer S1-34M

<400> 3
 gccaccagc tcattaaggc t 21
 <210> 4
 <211> 20
 <212> DNA
 <213> Artificial Sequence
 <220>
 <223> Primer A2-320M
 <400> 4
 tcattaatgg agcgtgggtg 20
 <210> 5
 <211> 20
 <212> DNA
 <213> Artificial Sequence
 <220>
 <223> Primer S2-55M
 <400> 5
 cctggcatca ctactgctat 20
 <210> 6
 <211> 18
 <212> DNA
 <213> Artificial Sequence
 <220>
 <223> Primer C375
 <400> 6
 ctgaacatca cgcccaac 18
 <210> 7
 <211> 19
 <212> DNA
 <213> Artificial Sequence
 <220>
 <223> Primer A1-350
 <400> 7
 aggaagcagc ggtggacca 19
 <210> 8
 <211> 20
 <212> DNA
 <213> Artificial Sequence
 <220>
 <223> Primer S1-34
 <400> 8

gcccacagc ttattaaggc	20
<210> 9	
<211> 20	
<212> DNA	
<213> Artificial Sequence	
<220>	
<223> Primer A2-320	
<400> 9	
tcattttattg agcggggatg	20
<210> 10	
<211> 20	
<212> DNA	
<213> Artificial Sequence	
<220>	
<223> Primer S2-55	
<400> 10	
cctggcatca ctactgctat	20
<210> 11	
<211> 25	
<212> DNA	
<213> Artificial Sequence	
<220>	
<223> Primer M1PR6	
<400> 11	
ccatgttcca caccgtattc cagag	25
<210> 12	
<211> 25	
<212> DNA	
<213> Artificial Sequence	
<220>	
<223> Primer S4294M	
<400> 12	
gtgttctacg gggatgctta tgacg	25
<210> 13	
<211> 25	
<212> DNA	
<213> Artificial Sequence	
<220>	
<223> Primer M1PF6	
<400> 13	
gactcagtat tctctgctgc cgtgg	25

<210> 14
 <211> 25
 <212> DNA
 <213> Artificial Sequence

 <220>
 <223> Primer A4556

 <400> 14
 ggctcaccag aatgcttctt ccaga 25

 <210> 15
 <211> 342
 <212> DNA
 <213> Hepatitis B Virus

 <220>
 <223> Clone: USP-15

 <400> 15
 gcccatcagt ttattaaggc tcttggcatt actactgcc ttagcaggc tgctctggct 60
 gcggccaatt ctgccttggc gaatgctgtg gtggttcggc cgtttttatc tcgcgtgcaa 120
 accgagattc ttattaattt gatgcaaccc cggcagttgg ttttcgccc tgaggtactt 180
 tggaatcacc ctatccagcg ggttatacat aatgaattag aacagtactg ccgggctcgg 240
 gctgggtcgtt gcttggaggt tggagctcac ccaagatcca ttaatgacaa cccaacggt 300
 ctgcacgggt gtttccttag accggtcggg cgtgatgttc ag 342

 <210> 16
 <211> 20
 <212> DNA
 <213> Artificial Sequence

 <220>
 <223> Primer PA2-5560

 <400> 16
 taggttatac tgccggcgca 20

 <210> 17
 <211> 20
 <212> DNA
 <213> Artificial Sequence

 <220>
 <223> Primer S1-5287

 <400> 17
 ttctcagccc ttcgcaatcc 20

 <210> 18
 <211> 20
 <212> DNA
 <213> Artificial Sequence

 <220>
 <223> Primer S2-5310

<400> 18
 atattcatcc aaccaacccc 20

 <210> 19
 <211> 251
 <212> DNA
 <213> Hepatitis E Virus

 <220>
 <223> Clone: b421

 <400> 19
 atattcatcc aaccaacccc ttgcgcgcg atgtcgtttc acaaccggg gctggaactc 60
 gccctcgaca gccgccccgc cccctcggtt ccgcttggcg tgaccagtcc cagcgccct 120
 ccgttgccc ccgtcgtcga tctaccocag ctggggctgc gccgctaact gccatatcac 180
 cagccctga tacagctect gtacctgatg ttgactcacg tgggtgtatt ttgcgcggc 240
 agtataacct a 251

 <210> 20
 <211> 20
 <212> DNA
 <213> Artificial Sequence

 <220>
 <223> Primer US4.2-69S/20

 <400> 20
 ttccgcttgg cgtgaccagt 20

 <210> 21
 <211> 21
 <212> DNA
 <213> Artificial Sequence

 <220>
 <223> Primer US4.4/144s

 <400> 21
 gctaactgcc atatcaccag c 21

 <210> 22
 <211> 20
 <212> DNA
 <213> Artificial Sequence

 <220>
 <223> Primer M6417a

 <400> 22
 cccttatcct gctgagcatt 20

 <210> 23
 <211> 24
 <212> DNA
 <213> Artificial Sequence

 <220>

<223> Primer M6371a

<400> 23
 ttggctcgcc attggctgag acaa 24

<210> 24
 <211> 899
 <212> DNA
 <213> Hepatitis E Virus

<220>
 <223> Clone: df-orf2/3

<400> 24
 gctaactgcc atatcaccag cccctgatac agctcctgta cctgatgttg actcacgtgg 60
 tgctatcttg cgccggcagt acaatttgtc taogtccccg cttacatcat ctggttcttc 120
 tggtaactaat ctgggtctct atgctgcccc gctgaaccct ctcttgccctc ttcaggatgg 180
 caccaacact catattatgg ctactgaggg atctaattac gccagtatc ggggtgttgg 240
 ggctacgatt cgttatcgcc cgttggtgcc aaatgctgtt ggtggttatg ctatctctat 300
 ttctttcttg cctcaaacta caactacccc tacttctgtt gacatgaatt ctatcacttc 360
 tactgatgtc aggatcttgg tccagcccg tatagcctcc gagttagtca tccctagtga 420
 acgccttcac taccgcaacc aaggctggcg ctctgttgag accacgggtg tggccgaaga 480
 ggaggctaac tccgtctggt taatgctttg tattcatggc tccctgtta actcctacac 540
 taatacacct tacacgggtg cattggggct tcttgatttt gcattagaac ttgaatttag 600
 aaatttgaca cccgggaaca ctaacaccgg tgtttcccg tatactagca cagcccgcca 660
 ccggctgcgc cgcggtgctg atgggaccgc tgagctcacc accacagcag ccacacgctt 720
 catgaaggat ttgcatctta ctggtacgaa cggcggttgg gaggtgggtc gtggtattgc 780
 cctgaactctg tttaatcttg ctgatacgtc tcttggtggt ttacgacag aattgatttc 840
 gtcggtggg ggtcaactgt tttactcccg ccctgttgtc tcagccaatg gogagccaa 899

<210> 25
 <211> 20
 <212> DNA
 <213> Artificial Sequence

<220>
 <223> Primer USP 3s/20

<400> 25
 tggcattact actgccattg 20

<210> 26
 <211> 20
 <212> DNA
 <213> Artificial Sequence

<220>
 <223> Primer M902A

<400> 26
 atcgatcgga catagacctc 20

<210> 27
 <211> 846
 <212> DNA
 <213> Hepatitis E Virus

<220>

<223> Clone: df-orf1

<400> 27

tgccattact	actgccattg	agcaggctgc	tctggctgcg	gccaatctctg	ccttgggcga	60
tgctgtggtg	gttcggccgt	ttttatctcg	cgtgcaaacc	gagattctta	ttaatttgat	120
gcaaccccg	cagttgggtt	tccgccctga	ggtacttttg	aatcaccccta	tccagcgggt	180
tatacataat	gaattagaac	agtactgccg	ggctcgggct	ggtcgttgct	tggagggttg	240
agctcaccga	agatccatta	atgacaacc	caacgtttctg	catcgggtgt	tccttagacc	300
ggttggccga	gatgttcagc	gctgggtactc	tgccccacc	cgcggccctg	cggctaattg	360
cgcgcgtcc	gcgttgcgtg	gtctcccc	cgtgaccgc	acttactgct	ttgatggatt	420
ctccggtgt	gcttttgctg	cagagaccgg	tgtggctctt	tactctctgc	atgaccttg	480
gccagctgat	gttgacagag	ctatggcccg	ccacgggatr	acacgcttgt	atgccgcact	540
gcaccttccc	cctgaggtgc	tgctaccacc	cggcacctac	cacacaacct	cgtatctcct	600
gattcacgac	ggcgaccg	ctgttgtaac	ttacgagggc	gatactagt	cgggctataa	660
tcattgatgtc	tccatacttc	gtgcgtggat	cgtactaca	aaaatagttg	gtgatcatcc	720
gttggtcata	gagcgtgtgc	gggccattgg	atgtcatttt	gtgttgctgc	tcaccgcagc	780
ccttgagcgc	tcaccatgc	cttatgttcc	ttaccctcgt	tcaacggagg	tctatgtccg	840
atcgat						846

<210> 28

<211> 23

<212> DNA

<213> Artificial Sequence

<220>

<223> Primer 3750s

<400> 28

cttccatcag	ttggctgagg	agc	23
------------	------------	-----	----

<210> 29

<211> 23

<212> DNA

<213> Artificial Sequence

<220>

<223> Primer 3900a

<400> 29

gccatgcggc	agtgacacat	gtc	23
------------	------------	-----	----

<210> 30

<211> 168

<212> DNA

<213> Hepatitis E Virus

<220>

<223> Clone: HEV 167

<400> 30

cttccatcag	ttggctgagg	agctgggcca	tcgcccggcc	cctgtcgcgc	ccgtcttgcc	60
cccttgccct	gagcttgagc	agggcctgct	ctacatgcca	caggagctca	ctgtgtccga	120
tagtgtgttg	gtttttgagc	ttacggacat	tgtgcactgc	cgcattggc		168

<210> 31

<211> 25

```

<212> DNA
<213> Artificial Sequence

<220>
<223> Primer 5000s

<400> 31
ctcgttcata acctgattgg catgc                25

<210> 32
<211> 23
<212> DNA
<213> Artificial Sequence

<220>
<223> Primer uf-orf2/3 a3

<400> 32
ggactgggtca cgccaagcgg aac                23

<210> 33
<211> 424
<212> DNA
<213> Hepatitis E Virus

<220>
<223> Clone: HEV 426

<400> 33
ctcgttcata acctgattgg catgctgcag accatcgccg atggcaaggc ccactttaca    60
gagactatta aacctgtact tgatctcaca aattccatca tacagcgggt ggaatgaata    120
acatgtcttt tgcctcgccc atgggatcac catgcgccct agggctgttc tgttggtgtt    180
cctcatgttt ctgcctatgc tgcccgcgcc accggccggg cagccgtctg gccgtcgccg    240
tgggcggcgc agcggcggtg ccggcggtgg tttctggagt gacagggttg attctcagcc    300
cttcgccctc ccttatattc atccaaccaa ccccttcgcc gccgatgtcg tttcacaacc    360
cggggctgga actcgccctc gacagccgcc ccgccccctc ggttcgctt ggcgtgacca    420
gtcc                                           424

<210> 34
<211> 24
<212> DNA
<213> Artificial Sequence

<220>
<223> Primer 167-s1

<400> 34
tctacatgcc acaggagctc actg                24

<210> 35
<211> 27
<212> DNA
<213> Artificial Sequence

<220>
<223> Primer 426-a3

```

<400> 35	
gatggaattt gtgagatcaa gtacagg	27
<210> 36	
<211> 25	
<212> DNA	
<213> Artificial Sequence	
<220>	
<223> Primer 167-s2	
<400> 36	
ctcactgtgt ccgatagtgt gttgg	25
<210> 37	
<211> 23	
<212> DNA	
<213> Artificial Sequence	
<220>	
<223> Primer 426-a4	
<400> 37	
ccttgccatc ggcgatggtc tgc	23
<210> 38	
<211> 1186	
<212> DNA	
<213> Hepatitis E Virus	
<220>	
<223> Clone: HEV 1186	
<400> 38	
ctcactgtgt ccgatagtgt gttggttttt gagcttacgg atatagttca ttgccgcatg	60
gcgcgtccaa gccagcgaaa ggctgtttctc tcaacacttg tggggaggta tggccgtagg	120
acgaaactat atgaggcggc gcattcagat gttcgtgagt ccctagctag gttcatccct	180
actatcgggc ctgttcaggc taccacatgt gagttgtatg agttggttga ggctatgggtg	240
gagaaaaggtc aggacggctc tgcagtctta gagcttgatc tttgtaatcg tgatgtctcg	300
cgcatcacat ttttccaaaa agwctgcaac aagtttataa ctggtgagac catcgccac	360
ggcaagggtg gccagggtat atcggcctgg agtaagacct tctgcgctct gttcgcccg	420
tggttccgcg ccattgaaaa agaaatattg gccctgctcc cgcctaatat cttttatggc	480
gacgcttatg aggagtcagt ttttgccgcc gctgtgtccg gggcggggtc atgtatggta	540
tttgaataatg acttttcaga gtttgacagt acccagaata atttctctct tggccttgag	600
tgtgtgggta tggaggagtg cggcatgcct caatggctaa ttaggttgta ccactcgggt	660
cggtctgcct ggattctgca ggcgcgaag gagtctctta agggtttctg gaagaagcat	720
tctggtgagc ctggtaccct tctttggaat accgtctgga atatggcgat tatagcacat	780
tgctatgagt tccgtgactt tccgtgtgct gcccttaagg gtgatgatc ggtggtcctc	840
tgtagtgact accgacagag ccgcaatgca gctgccttaa ttgctggctg tgggctcaaa	900
ttgaagggtg attaccgccc tatcgggctg tatgctgggg tgggtggggc ccccggttg	960
gggacactgc ccgatgtggt gcgttttgcct ggtcgggtgt ctgaaaagaa ttggggcccc	1020
ggcccggaac gtgctgagca gctgcgtctt gctgtctgcg acttccttcg aggggtgacg	1080
aatgttgccg aggtctgtgt tgatgttggt tccgtgtct atggagtcag ccccgggctc	1140
gtacataacc ttattggcat gctgcagacc atcgccgatg gcaagg	1186
<210> 39	
<211> 25	

<212> DNA
 <213> Artificial Sequence

 <220>
 <223> Primer orf1-s2

 <400> 39
 tcacccatgc ottatgttcc ttacc 25

 <210> 40
 <211> 22
 <212> DNA
 <213> Artificial Sequence

 <220>
 <223> Primer 1300a

 <400> 40
 ggccggcctgg gatgtaata ca 22

 <210> 41
 <211> 460
 <212> DNA
 <213> Hepatitis E Virus

 <220>
 <223> Clone: HEV 459

 <400> 41
 tcacccatgc ottatgttcc ttaccctcgt tcaacggagg tgtatgtccg gtccatattt 60
 ggccctggcg gctcccatc ottgtttccg tcagcctgct ctactaaatc tactttccat 120
 gctgtcccg tgcatatctg ggatcggtc atgctctttg gtgccaccct ggacgatcag 180
 gcgttttgcgt gttcacggct catgacttac ctccgtggta ttagttacaa ggtcactgtc 240
 ggcgcgcttg tcgctaata ggggtggaac gcctctgaag acgctcttac tgcartgac 300
 actgcagctt atttgactat ttgccatcag cgttatctcc gcaccaggg gatatccaag 360
 ggcacgcgc ggttgggggt tgagcacgcc cagaaattta tcacaagact ctacagttgg 420
 ctatttgaga agtctggcgg tgattacatc ccaggccgcc 460

 <210> 42
 <211> 26
 <212> DNA
 <213> Artificial Sequence

 <220>
 <223> Primer 459-s2

 <400> 42
 cagaaattta tcacaagact ctacag 26

 <210> 43
 <211> 23
 <212> DNA
 <213> Artificial Sequence

 <220>
 <223> Primer 1450a

<400> 43
 aacactcctg accgagccac ttc 23

 <210> 44
 <211> 235
 <212> DNA
 <213> Hepatitis E Virus

 <220>
 <223> Clone: HBV 216

 <400> 44
 cagaaattta tcacaagact ctacagttgg ctatttgaga agtctggccg tgattatata 60
 cccggccgcc agcttcagtt ctatgcacag tgccgacggt ggctatctgc aggcttccac 120
 ctagacccca ggggtacttgt ttttgatgag tcagtaccat gccgctgtag gacgtttttg 180
 aagaaagttg cgggtaaatt ctgctgtttt atgaagtggc tcggtcagga gtgtt 235

 <210> 45
 <211> 26
 <212> DNA
 <213> Hepatitis E Virus

 <220>
 <223> us1 gap-s1

 <400> 45
 tatagatata acaggttcac ccagcg 26

 <210> 46
 <211> 24
 <212> DNA
 <213> Hepatitis E Virus

 <220>
 <223> us1 gap-a0.5

 <400> 46
 gctgcaagac cctcacgcat gatg 24

 <210> 47
 <211> 23
 <212> DNA
 <213> Hepatitis E Virus

 <220>
 <223> us1 gap-s2

 <400> 47
 cggattatgg ttacaccctg agg 23

 <210> 48
 <211> 25
 <212> DNA
 <213> Hepatitis E Virus

 <220>
 <223> us1 gap-a1

<400> 48		
attcagttgg	gtaaaaacgct tctgg	25
<210> 49		
<211> 545		
<212> DNA		
<213> Hepatitis E Virus		
<220>		
<223> us1-gap		
<400> 49		
cggtattatgg	ttacaccctg agggggttgc	60
tatctgggag	tctgcgaacc ccttttgcgg	120
aacatctggc	ttttctagtg atttctcccc	180
taccccgggg	ctgccccatt ctacccacc	240
ctcagaggag	tttcagggtg atgcagcacc	300
gcccgggtccc	gttggtctta ccccccccc	360
cccgcccttc	cgtaaccgtc gtctctctta	420
agggtcactg	tttgaatcag actgtgactg	480
tcccggaggt	ggcctctgcc atgcctttta	540
tgaat	ccaacgtttt ccagaagcgt	545
<210> 50		
<211> 24		
<212> DNA		
<213> Hepatitis E Virus		
<220>		
<223> us1-2600s		
<400> 50		
gtgctcacca	taactgagga cacg	24
<210> 51		
<211> 24		
<212> DNA		
<213> Hepatitis E Virus		
<220>		
<223> us1-2600a		
<400> 51		
cgctgcata	gtaacagcaa cagg	24
<210> 52		
<211> 344		
<212> DNA		
<213> Hepatitis E Virus		
<220>		
<223> us1-344		
<400> 52		
gtgctcacca	taactgagga caggcccggt	60
gctacagagg	tggcccggtg tgcacccatca	120

```

cagtttaccg cccgggtccc gggctcgggc aagtcaaggt ccatacaaca gggagatgtc 180
gatgtgggtg ttgtgcccac ccgggagctt cgtaatagtt ggccgcgcgc gggttttgcg 240
gccttcacac cccacacagc ggcccgtgtt actatcggcc gccgcgttgt gattgatgag 300
gctccatctc tcccgccaca cctgttgctg ttacatatgc agcg 344

```

```

<210> 53
<211> 23
<212> DNA
<213> Hepatitis E Virus

```

```

<220>
<223> us1 3200s

```

```

<400> 53
gccgatgtgt gcgagctcat acg 23

```

```

<210> 54
<211> 25
<212> DNA
<213> Hepatitis E Virus

```

```

<220>
<223> us1 3200a

```

```

<400> 54
atgattgtgg tctctgtgaa ggtgg 25

```

```

<210> 55
<211> 194
<212> DNA
<213> Hepatitis E Virus

```

```

<220>
<223> us1-194

```

```

<400> 55
gccgatgtgt gcgagctcat acgcggagcc taccctaaaa tccagaccac gagecgtgtg 60
ctacgggtccc tgttttggaa tgaaccggcc attggccaga agttggttyt cacgcaggcg 120
gcaaaggctg ctaaccctgg tgcgattacg gtccacgaag ctcaggggtgc caccttcaca 180
gagaccacaa tcat 194

```

```

<210> 56
<211> 23
<212> DNA
<213> Hepatitis E Virus

```

```

<220>
<223> HEV216-s1

```

```

<400> 56
cagtaccatg ccgctgtagg acg 23

```

```

<210> 57
<211> 26
<212> DNA
<213> Hepatitis E Virus

```

<220>
 <223> us2-733a1

 <400> 57
 ccattagatg aaatctttac ctgcag 26

 <210> 58
 <211> 26
 <212> DNA
 <213> Hepatitis E Virus

 <220>
 <223> HEV216-s2

 <400> 58
 gtaggacgtt tttgaagaaa gttgcg 26

 <210> 59
 <211> 24
 <212> DNA
 <213> Hepatitis E Virus

 <220>
 <223> us2-733a2

 <400> 59
 ggtgagctca taagtgaggc tgtg 24

 <210> 60
 <211> 464
 <212> DNA
 <213> Hepatitis E Virus

 <220>
 <223> us1-733wb

 <400> 60
 gtaggacgtt tttgaagaaa gttgcgggta aattctgctg ttttatgcgg tggctcgggc 60
 aggagtgtac ctgcttcttg gagccggccg aggggttagt cggcgatcat ggccatgaca 120
 acgaggccta tgagggttct gaggtcgacc cggctgaacc tgcacatctt gatgtttctg 180
 ggacttacgc cgtccacggg caccagcttg aggccctcta tagggcactt aatgtccac 240
 aagatattgc cgtcgcagct tcccgactaa cggcaactgt tgagctcgtt gcaagtccag 300
 accgcttaga gtgcgcgacc gtgctcggta ataagacctt ccggacgacg gtggtcgacg 360
 gcgcccattc agaggcgaat ggccctgagc agtatgtctt atcatttgac gcctcccgtc 420
 agtctatggg ggccgggtcg cacagcctca cttatgagct cacc 464

 <210> 61
 <211> 24
 <212> DNA
 <213> Hepatitis E Virus

 <220>
 <223> us1 733s1

 <400> 61
 ttgagctcgt tgcaagtcca gacc 24

```

<210> 62
<211> 22
<212> DNA
<213> Hepatitis E Virus

<220>
<223> us2851-r2

<400> 62
ccagagggttg accaggttcg gg                22

<210> 63
<211> 24
<212> DNA
<213> Hepatitis E Virus

<220>
<223> us1 733s2

<400> 63
ccgtgctcgg taataagacc ttcc                24

<210> 64
<211> 433
<212> DNA
<213> Hepatitis E Virus

<220>
<223> us1-432

<400> 64
ccgtgctcgg taataagacc ttccggacga cggtaggtcga cggcgcccat ctagaggcga        60
atggccctga gcagtatgtc ttatcatttg acgcctcccg tcagtctatg ggggcccgggt        120
cgcatagcct cacttatgag ctaccccctg ctggtttgca ggtaggatt tcactctaatg        180
gtctggattg cactgtctaca ttcccccccg gtggagcccc tagcgtctgcg cccggggagg        240
tggcagcctt ttgcagtgcc ctttatagat ataacagggt caccagcgg cactcgtga        300
ctggcggatt atggttacac cctgaggggt tgctgggtat ttccccccct ttctcccctg        360
ggcatatctg ggagtctgcg aaccctttt gggggagggg gactttgtat acccgaacct        420
ggtcaacctc tgg                                433

<210> 65
<211> 26
<212> DNA
<213> Hepatitis E Virus

<220>
<223> us2851-f1

<400> 65
gactgtgatt ggtagtcaa tgcctc                26

<210> 66
<211> 24
<212> DNA
<213> Hepatitis E Virus

<220>

```

```

<223> us1 430-a1

<400> 66
cgtgtcctca gttatggtga gcac                24

<210> 67
<211> 26
<212> DNA
<213> Hepatitis E Virus

<220>
<223> us1 430-a2

<400> 67
tattagcctc aaaccaatth gcagcg                26

<210> 68
<211> 382
<212> DNA
<213> Hepatitis E Virus

<220>
<223> us1-382

<400> 68
gactgtgatt ggtagtcaa tgctcaaac ccgggccatc gtcccgagg tggcctctgc        60
catgcctttt accaacgttt tccagaagcg ttttaccaa ctgaattcat catgcgtgag        120
ggtcttgtag catacacctt gaccccgcg cctatcatc atgcagtcgc tcccgattat        180
agggttgagc agaaccggaa gaggtctgag gcagcgtacc gtgaaacttg ttcccgctgt        240
ggcaccgctg cctaccgct tttgggttcg ggtatatacc aggtccctgt tagcctcagt        300
tttgatgcct gggaacgtaa tcaccgccc gcgatgagc tttacttgac cgagcccgct        360
gcaaattggt ttgaggctaa ta                382

<210> 69
<211> 22
<212> DNA
<213> Hepatitis E Virus

<220>
<223> us2-579-s1

<400> 69
cagaccacga gccgtgtgct ac                22

<210> 70
<211> 25
<212> DNA
<213> Hepatitis E Virus

<220>
<223> JE hev167-a1

<400> 70
ccaacacact atcgacaca gtgag                25

<210> 71
<211> 22

```

<212> DNA
 <213> Hepatitis E Virus

 <220>
 <223> us2-579-s2

 <400> 71
 gctgctaagg ctgccaaccc tg 22

 <210> 72
 <211> 24
 <212> DNA
 <213> Hepatitis E Virus

 <220>
 <223> JE hev167-a2

 <400> 72
 cagtgaagctc ctgtggcatg taga 24

 <210> 73
 <211> 451
 <212> DNA
 <213> Hepatitis E Virus

 <220>
 <223> us1-579wb

 <400> 73
 gctgctaagg ctgccaaccc tgggtgcgatt acgggtccacg aagctcaggg tgccaccttc 60
 acagagacca caatcatagc caaggccgac gccagggggcc ttatccagtc atccccgggt 120
 catgctatag ttgcacttac tcgccacact gagaagtgtg ttatcctgga tgccccgggc 180
 ctgcttcgtg aggtcggcat ttccgcatgtg attgtcaaca actttttcct tgctgggtggc 240
 gaggtcggcc rccaccgccc ttctgtgata cctcgcggta accctgatca aaacctcggg 300
 actttacagg ccttcccgcc gtccgtgtcaa attagtgtt accatcagtt ggctgaggaa 360
 ctgggccatc gcccgggccc tgtcgcgcgc gtcttgcccc cttgccctga gcttgagcag 420
 ggctgtctct acatgccaca ggagctcact g 451

 <210> 74
 <211> 24
 <212> DNA
 <213> Hepatitis E Virus

 <220>
 <223> us2-430s1

 <400> 74
 ggtatatacc aggtccctgt tagc 24

 <210> 75
 <211> 22
 <212> DNA
 <213> Hepatitis E Virus

 <220>
 <223> us2-482-a1

<400> 75
 ccgctgtgtg aggtgtgaag gc 22

 <210> 76
 <211> 24
 <212> DNA
 <213> Hepatitis E Virus

 <220>
 <223> us2-430s2

 <400> 76
 gttagcctca gttttgatgc ctgg 24

 <210> 77
 <211> 23
 <212> DNA
 <213> Hepatitis E Virus

 <220>
 <223> us2-482-a2

 <400> 77
 gacgccagct gttacggagc tcc 23

 <210> 78
 <211> 334
 <212> DNA
 <213> Hepatitis E Virus

 <220>
 <223> us1-430wb

 <400> 78
 gttagcctca gttttgatgc ctgggaacgt aatcacccgc ccggcgatga gctttacttg 60
 accgagcccg ctgcaaattg gtttgaggct aataagccgg cgcagccggg gctcaccata 120
 actgaggaca cggcccgtag ggccaacctg gcattggaga ttgatgccgc tacagaggtc 180
 ggccgtgctt gtgccgggtg caccatcagc cctggcattg tgcactatca gtttacggcc 240
 ggggtccccg gctcgggcaa gtcaagggtc atacaacagg gagatgtcga tgtgggtggt 300
 gtgcccaccc gggagctccg taacagctgg cgtc 334

 <210> 79
 <211> 23
 <212> DNA
 <213> Hepatitis E Virus

 <220>
 <223> us2-482-s1

 <400> 79
 gatgtcgatg tgggtggtgt gcc 23

 <210> 80
 <211> 23
 <212> DNA
 <213> Hepatitis E Virus

<220>
 <223> JE us2-579-a1

 <400> 80
 gtaatcgac cagggttggc agc 23

 <210> 81
 <211> 23
 <212> DNA
 <213> Hepatitis E Virus

 <220>
 <223> us2-482-s2

 <400> 81
 ggagctccgt aacagctggc gtc 23

 <210> 82
 <211> 22
 <212> DNA
 <213> Hepatitis E Virus

 <220>
 <223> JE us2-579-a2

 <400> 82
 cagggttggc agccttagca gc 22

 <210> 83
 <211> 413
 <212> DNA
 <213> Hepatitis E Virus

 <220>
 <223> us1-482wb

 <400> 83
 ggagctccgt aacagctggc gtcgccggg ttttgcggc ttcacacccc acacagcggc 60
 ccgtgttact atcgcccgcc gcgttgtgat tgatgaggct ccatctctcc cgccacacct 120
 gttgctgtta catatgcagc gggcctcctc ggtccatctc ctccgtgacc caaatcagat 180
 ccctgctatt gattttgagc acgccggcct ggtccctgcy atccgtcccg agcttgogcc 240
 aacgagctgg tggcgcgtta cacaccgttg cccggccgat gtgtgcgagc tcatacgcgg 300
 agcctaccct aaaatccaga ccacgagccg tgtgctacgg tccctgtttt ggaatgaacc 360
 ggccattggc cagaagttgg ttytcacgca gctgctaag gctgccaacc ctg 413

 <210> 84
 <211> 37
 <212> DNA
 <213> Artificial Sequence

 <220>
 <223> Oligo dT Adapter Primer

 <400> 84
 ggccacgcgt cgactagtagc tttttttttt ttttttt 37

 <210> 85

<211> 20
 <212> DNA
 <213> Artificial Sequence

 <220>
 <223> AUAP Primer

 <400> 85
 ggccacgcgt cgactagtac 20

 <210> 86
 <211> 22
 <212> DNA
 <213> Artificial Sequence

 <220>
 <223> Primer df-orf3-s1

 <400> 86
 gcgttggtga ggtgggtcgt gg 22

 <210> 87
 <211> 24
 <212> DNA
 <213> Artificial Sequence

 <220>
 <223> Primer df-orf3-s2

 <400> 87
 cgcttccttg tggtttaccg acag 24

 <210> 88
 <211> 960
 <212> DNA
 <213> Hepatitis E Virus

 <220>
 <223> Clone: HEV 3p RACE

 <400> 88
 cgcttccttg tggtttaccg acagaattga ttctgtcggc tgggggtcaa ctgttttact 60
 ccgcacctgt tgtctcggcc aatggcgagc caacagtaaa gttatacaca tctgttgaga 120
 atgcgcagca agacaagggc atcaccattc cacacgacat agatttaggt gactcccgtg 180
 tggttatcca ggattatgat aaccagcacg aacaagatcg acctaccccg tcacctgccc 240
 cctcccgccc tttctcagtt ctctgtgcca atgatgtttt gtggctctct ctactgccc 300
 ctgagtacgr ccagaccacg tatgggtcgt ccaccaaccc tatgtatgtc tctgatacag 360
 tcacgcttgt taatgtagcc actgggtgctc aggtctgttc ccgctctctt gactgggtcta 420
 aagttactct ggatgggtcgc cctcttacta ccattcagca gtattctaag aaattttatg 480
 ttctcccgtc tcgsgggaag ctgtcctttt gggaggctgg tacgaccaag gccggctacc 540
 cgtataatta taataccact gctagtgaac aaattttgat tgagaacgcg gccggtcacc 600
 gtgtcgccat ttctaactat accactagtt tgggtgcccg ccctacctcg atytctgcgg 660
 tcggtgtact agctccacat tcggcccttg ctgttctcga ggatactgtt gattatcctg 720
 ctctgtccca tacttttgat gatttctgcc cggagtgtcg cacccttggc ctgcagggtt 780
 gtgcattcca atctactatt gctgaacttc agcgtcttaa aatgaaggta ggtaaaaccc 840
 gggagtctta attaattcct tttgtgcccc ctctgcagtt ctctttggct ttatttctca 900
 tttctgcttt ccgcgctncc ctggaaaaaa aaaaaaaaaa gtactagtgc acgcgtggcc 960

<210> 89
 <211> 7202
 <212> DNA
 <213> Hepatitis E Virus

<220>
 <223> us1ful1

<400> 89
 cctggcatta ctactgccat tgagcaggct gctctggctg cggccaattc tgccttggcg 60
 aatgctgttg tggttcggcc gtttttatct cgcgtgcaaa ccgagattct tattaatttg 120
 atgcaacccc ggcagttggt tttccgccct gaggtacttt ggaatcacc cccagaggg 180
 gttatacata atgaattaga acagtactgc cgggctcggg ctggtcgttg cttggagggt 240
 ggagctcacc caagatccat taatgacaac cccaacgttc tgcacgttg tttccttaga 300
 ccggttggcc gagatgttca gcgctgttac tctgccccca cccgcgcccc tgcggctaatt 360
 tgcgcgcgct ccgcgttgcg tgggtctccc cccgctgacc gcacttactg ctttgatgga 420
 ttctcccggt gtgcttttgc tgcagagacc ggtgtggctc tttactctct gcacgtactt 480
 tggccagctg atgttgacaga ggctatggcc cgcacgggga trcacgcgtt gtatgcgcga 540
 ctgcaccttc cccctgaggt gctgctacca cccggcacct accacacaac ctcgatatct 600
 ctgattcacg acggcgaccg cgcgtgtgta acttacgagg gcgatactag tgcgggctat 660
 aatcatgatg tctccatact tcgtgcgttg atccgtacta caaaaatagt tgggtgatcat 720
 ccgttgggtca tagagcgtgt gcggggccatt ggatgtcatt ttgtgttgct gctcacgcga 780
 gcccctgagc cgtcacccat gcttatgttt ccttaccctc gttcaacgga ggtgtatgtc 840
 cggctccatg ttggccctgg cggctcccca tccttgtttc cgtcagcctg ctctactaaa 900
 tctactttcc atgctgtccc ggtgcatacc tgggatcggc tcatgctctt tgggtgccacc 960
 ctggacgatc aggcgttttg ctgttcacgg ctcatgactt acctccgttg tattagttac 1020
 aaggtcactg tcggcgcgct tgcgtgtaat gaggggtgga acgcctctga agacgctctt 1080
 actgcartga tcaactgcagc ttatttgact atttgccatc agcgttatct ccgcacccag 1140
 gcgatatcca agggcatgcg ccggttgggg gttgagcacg cccagaaatt tatcacaaga 1200
 ctctacagtt ggctatttga gaagtctggc cgtgattata tccccggccg ccagcttcag 1260
 ttctatgcac agtgccgacg gtggctatct gcaggcttcc acctagaccc cagggtactt 1320
 gtttttgatg agtcagtacc atgcgcgtgt aggcgtttt tgaagaaagt tgcgggtaaa 1380
 ttctgctggt ttatgcgggt gctcgggcag gactgtacct gcttcttggg gccggccgag 1440
 ggttttagtc gcgatcatgg ccacgacaac gaggcctatg agggttctga ggtcgacccg 1500
 gctgaacctg cacatcttga tgtttctggg acttacgcgg tccacgggca ccagcttgag 1560
 gccctctata gggcacttaa tgtcccaaaa gatattgccc ctccgagctt ccgactaacg 1620
 gcaactgttg agctcgttgc aagtccagac cgttagaggt gccgcacgtt gctcggtaat 1680
 aagaccttcc ggacgacggt ggtcgacggc gccatcttag aggcgaatgg ccctgagcag 1740
 tatgtcttat catttgacgc ctcccgtcag tctatggggg ccgggtcgca tagcctcact 1800
 tatgagctca cccctgctgg tttgcagggt aggatctcat ctaatggtct ggattgcact 1860
 gctacattcc cccccggtgg agccccatgc gctgcgcccg gggagggtggc agccttttgc 1920
 agtgccttct atagatataa cagggttcacc cagcggcact cgtgactgg cggattatgg 1980
 ttacacctcg aggggttgct ggttatttct ccccttctct cccctgggca tatctgggag 2040
 tctgcgaacc ccttttgcgg ggaggggact ttgtataccc gaacttgggt aacatctggc 2100
 ttttctagtg atttctcccc ccctgaagcg gccgctctg ctatggctgc taccgccggg 2160
 ctgccccatt ctacccacc tgttagcgat atttgggtgc taccaccgpc ctacagaggag 2220
 ttocagggtg atgcagcacc tgtgccccct gccctgacc ctgctggatt gcccggtccc 2280
 gttgtgctta ccccccccc cctccccct gtgcataagc catcaatacc cccgccttcc 2340
 cgtaaccgtc gtctcctcta tacctatcct gacggcgcta aggtgtatgc agggctcactg 2400
 tttgaatcag actgtgactg gctggttaat gcctcaaac ccggccatcg tcccgagggt 2460
 ggctctgccc atgcttttta ccaacgtttt ccagaagcgt tttacccaac tgaattcact 2520
 atgctgaggg gtcttgacgc atacaccttg accccgcgcc ctatcattca tgcagtgcct 2580
 cccgattata ggggttgagca gaaccggaag aggcctgagg cagcgtaacc tgaacttgtt 2640
 tcccgctcgt gcaccgctgc ctacccgctt ttgggttcgg gtatatacca ggtccctggt 2700
 agcctcagtt ttgatgcctg ggaacgtaat caccgccccg gcgatgagct ttacttgacc 2760
 jagcccgctg caaattgggt tgaggccta aagccggcgc agccgggtgct caccataact 2820

gaggacaagg	cccgtagggc	caacctggca	ttggagattg	atgcccgtac	agaggtcggc	2880
cgtgcttgtg	ccgggttgac	catcagccct	ggcattgtgc	actatcagtt	tacggccggg	2940
gtccccggct	cgggcaagtc	aaggccata	caacagggag	atgtcgatgt	ggtgggtgtg	3000
cccacccggg	agcttcgtaa	tagttggcgc	cgccgggggt	ttggcgccct	cacacccac	3060
acagcgggcc	gtgttactat	cggccggcgc	gttgtgattg	atgaggtctc	atctctcccg	3120
ccacacctgt	tgctgttaca	tatgcagcgg	gcctcctcgg	tccatctcct	cggtagacca	3180
aatcagatcc	ctgctattga	ttttgagcac	gccggcctgg	tccttcgat	cggtcccgag	3240
cttgcgccaa	cgagctgggt	gcrcgttaca	caccgttgcc	cggccgatgt	gtgcgagctc	3300
atacggggag	cctaccctaa	aatccagacc	acgagccgtg	tgctacggtc	cctgtttttg	3360
aatgaaccgg	ccattggcca	gaagttgggt	ytacgcgagg	cggcaaaagg	tgctaaccct	3420
ggtgcgatta	cgggtccaaga	agctcagggg	gccaccttca	cagagaccac	aatcatagcc	3480
acggccgacg	ccaggggcct	tatccagtca	tcccgggctc	atgctatagt	tgcacttact	3540
cgccacactg	agaagtgtgt	tatcctggat	gccccgggcc	tgcttcgtga	ggtcggcatt	3600
tcggatgtga	ttgtcaacaa	ctttttcctt	gctggtggcg	aggtcggccr	ccaccgocct	3660
tctgtgatac	ctcgcggtaa	ccctgatcaa	aacctcggga	ctttacaggg	cttcccgcgg	3720
tcctgtcaaa	ttagtgtcta	ccatcagttg	gctgaggaac	tgggcatcgc	cccggccctt	3780
gtcgccgcgg	tcttgccccc	ttgccctgag	cttgagcagg	gcctgctcta	catgccacag	3840
gagctcactg	tgctcgatag	tggtttgggt	tttgagctta	cggatatagt	tcattgcccgc	3900
atggccgctc	caagccagcg	aaaggctggt	ctctcaacac	ttgtggggag	gtatggccgt	3960
aggaagaaac	tatatgaggg	ggcgcatcca	gatgttcgtg	agtccttagc	taggttcac	4020
cctaactatcg	ggcctgttca	ggctaccaca	tgtaggttgt	atgagttggg	tgaggctatg	4080
gtggagaaaag	gtcaggacgg	ctctgcagtc	ttagagcttg	atctttgtaa	tcgtgatgtc	4140
tcgcgcacatca	cattttttcca	aaaagwctgc	aacaagttta	caactggtga	gaccatcgcc	4200
cacggcaagg	ttggccaggg	tatatcgccc	tggagtaaga	ccttctgcgc	tctgttcggc	4260
ccgtggttcc	gcccattga	aaaagaaata	ttggccctgc	tcccgcctaa	tatcttttat	4320
ggcgacgctt	atgaggagtc	agtttttgcc	gccgctgtgt	ccggggcggg	gtcatgtatg	4380
gtatttgaaa	atgacttttc	agagtttgac	agtaaccaga	ataattttct	tcttgccctt	4440
gagtggtgtg	ttatggagga	gtgcggcatg	cctcaatggc	taattagggt	gtaccatctg	4500
gttcggtctg	cctggattct	gcaggcgccg	aaggagtctc	ttaagggttt	ctggaagaag	4560
cattctggtg	agcctggtag	ccttcttttg	aataccgtct	ggaatatggc	gattatagca	4620
cattgctatg	agttccgtga	ctttcgtgtt	gctgccttta	aggggtgatga	ttcgggtgtc	4680
ctctgtagtg	actaccgaca	gagccgcaat	gcagctgcct	taattgctgg	ctgtgggctc	4740
aaattgaagg	ttgattaccg	ccctatcggg	ctgtatgctg	gggtggtggt	ggccccgggt	4800
ttggggacac	tgcccgatgt	ggtgcgtttt	gctggctcgt	tgtctgaaaa	gaattggggc	4860
ccgggcccgg	aacgtgctga	gcagctcggt	cttgctgtct	gcgacttcc	tcgagggttg	4920
acgaatgttg	cgcaggtctg	tggtgatgtt	gtgtcccggt	tctatggagt	cagcccgggg	4980
ctcgtacata	accttatttg	catgctgcag	accatcgccg	atggcaaggc	ccactttaca	5040
gagactatta	aacctgtact	tgatctcaca	aattccatca	tacagcgggt	ggaatgaata	5100
acatgtcttt	tgcatcgccc	atgggatcac	catgcccct	agggctgttc	tggtgtgtgt	5160
cctcatgttt	ctgcctatgc	tgcccgcgcc	accggccggg	cagccgtctg	gccgtcgccg	5220
tgggcccggc	agcggcggtg	ccggcggtgg	ttcttgaggt	gacaggggtg	attctcagcc	5280
cttcgcccct	ccctatattc	atccaaccaa	cccttcgcgc	gccgatgtcg	tttcacaacc	5340
cggggcttga	actcgccctc	gacagccgcc	ccgcccctc	ggttcgcgtt	ggcgtgacca	5400
gtccaagcgc	ccctccgttg	ccccccgtcg	togatctacc	ccagctgggg	ctgcgcgct	5460
aactgccata	tcaccagccc	ctgatacagc	tcctgtacct	gatgttgact	cacgtggtgc	5520
tattttgctg	cggcagtaga	atttgtctac	gtccccgctt	acatcatctg	ttgcttctgg	5580
tactaatctg	gttctctatg	ctgccccgct	gaacctctc	ttgcctcttc	aggatggcac	5640
caacactcat	attatggcta	ctgaggcatc	taattacgcc	cagtatcggg	ttgttcgggc	5700
tacgattcgt	tatcgccgt	tgggtgcaaa	tgctgttggt	ggttatgcta	tctctatttc	5760
ttcttgccct	caaaactaca	ctacccctac	ttctgttgac	atgaattcta	tcacttctac	5820
tgatgtcagg	atcttggtcc	agcccgggtat	agcctccgag	ttagtcatcc	ctagtgaacg	5880
ccttcactac	cgcaaccaag	gctggcgctc	tggtgagacc	acgggtgtgg	ccgaagagga	5940
ggctacctcc	ggctcgtgta	tgctttgtat	tcattggctc	cctgttaact	cctacactaa	6000
tacacottac	accggtgcat	tggggcttct	tgatttttgc	ttagaacttg	aatttagaaa	6060
tttgacaccc	gggaacacta	acacccgtgt	ttcccgggtat	actagcacag	cccgccaccg	6120
gctgcgcgcg	ggtgctgatg	ggaccgctga	gctcaccacc	acagcagcca	cacgcttcat	6180
gaaggatttg	cattttactg	gtacgaacgg	cgttggtgag	gtgggtcgtg	gtattgcctt	6240

```

gactctgttt aatcttgctg atacgcttct tggtaggtta cgcacagaat tgatttcgtc 6300
ggctgggggt caactgtttt actccgccc tgtgtctctg gccaatggcg agccaacagt 6360
aaagttatac acatctgttg agaatgcgca gcaagacaag ggcatcacca ttccacacga 6420
catagattta ggtgactccc gtgtgggtat ccaggattat gataaccagc acgaacaaga 6480
tcgacctacc cgcgcacctg cccctcccgc ccctttctca gttcttcgtg ccaatgatgt 6540
tttgtggctc tctctcactg ccgctgagta cgrccagacc acgtatgggt cgtccaccaa 6600
ccctatgtat gtctctgata cagtcacgct tgttaatgta gccactgggt ctcaggctgt 6660
tgcccgctct cttgactggg cttaaagttac tctggatggg cgccctctta ctaccattca 6720
gcagtattct aagaaatttt atgttctccc gcttcgsggg aagctgtcct tttgggaggc 6780
tggtagcacc aaggccggct acccgataaa ttataatacc actgctagt accaaaattt 6840
gattgagaac gcggccggct accgtgtcgc catttctact tataccacta gtttgggtgc 6900
cggccctacc tcgatyctg cggtcggtgt actagctcca cattcggcc ttgctgttct 6960
cgaggatact gttgattatc ctgctcgtgc ccatactttt gatgatttct gcccgagtgc 7020
tcgcaccctt ggtctgcagg gttgtgcatt ccaatctact attgctgaac ttcagcgtct 7080
taaaatgaag gtaggtaaaa cccgggagtc ttaattaatt ccttttgtgc ccccttcgca 7140
gttctctttg gctttatttc tcatttctgc tttccgcgct ccctggaaaa aaaaaaaaaa 7200
aa

```

```

<210> 90
<211> 7202
<212> DNA
<213> Hepatitis E Virus

<220>
<223> us1ful1

<221> CDS
<222> (1)...(5094)
<223> Orf1

<221> CDS
<222> (5132)...(7111)
<223> Orf2

<223> Orf3 at positions 5094-5462

<223> Xaa = Unknown or Other at position 174
<223> Xaa = Unknown or Other at position 363
<223> Xaa = Unknown or Other at position 1088
<223> Xaa = Unknown or Other at position 1131
<223> Xaa = Unknown or Other at position 1217
<223> Xaa = Unknown or Other at position 1389
<223> Xaa = Unknown or Other at position 2179
<223> Xaa = Unknown or Other at position 2240
<223> Xaa = Unknown or Other at position 2293

<400> 90

```

cct ggc att act act gcc att gag cag gct gct ctg gct gcg gcc aat	48
Pro Gly Ile Thr Thr Ala Ile Glu Gln Ala Ala Leu Ala Ala Asn	
1 5 10 15	
tct gcc ttg gcg aat gct gtg gtg gtt cgg ccg ttt tta tct cgc gtg	96
Ser Ala Leu Ala Asn Ala Val Val Val Arg Pro Phe Leu Ser Arg Val	
20 25 30	
caa acc gag att ctt att aat ttg atg caa ccc cgg cag ttg gtt ttc	144
Gln Thr Glu Ile Leu Ile Asn Leu Met Gln Pro Arg Gln Leu Val Phe	
35 40 45	
cgc cct gag gta ctt tgg aat cac cct atc cag cgg gtt ata cat aat	192
Arg Pro Glu Val Leu Trp Asn His Pro Ile Gln Arg Val Ile His Asn	
50 55 60	
gaa tta gaa cag tac tgc cgg gct cgg gct ggt cgt tgc ttg gag gtt	240
Glu Leu Glu Gln Tyr Cys Arg Ala Arg Ala Gly Arg Cys Leu Glu Val	
65 70 75 80	
gga gct cac cca aga tcc att aat gac aac ccc aac gtt ctg cat cgg	288
Gly Ala His Pro Arg Ser Ile Asn Asp Asn Pro Asn Val Leu His Arg	
85 90 95	
tgt ttc ctt aga ccg gtt ggc cga gat gtt cag cgc tgg tac tct gcc	336
Cys Phe Leu Arg Pro Val Gly Arg Asp Val Gln Arg Trp Tyr Ser Ala	
100 105 110	
ccc acc cgc ggc cct gcg gct aat tgc cgc cgc tcc gcg ttg cgt ggt	384
Pro Thr Arg Gly Pro Ala Ala Asn Cys Arg Arg Ser Ala Leu Arg Gly	
115 120 125	
ctc ccc ccc gct gac cgc act tac tgc ttt gat gga ttc tcc cgt tgt	432
Leu Pro Pro Ala Asp Arg Thr Tyr Cys Phe Asp Gly Phe Ser Arg Cys	
130 135 140	
gct ttt gct gca gag acc ggt gtg gct ctt tac tct ctg cat gac ctt	480
Ala Phe Ala Ala Glu Thr Gly Val Ala Leu Tyr Ser Leu His Asp Leu	
145 150 155 160	
ttg cca gct gat gtt gca gag gct atg gcc cgc cac ggg atr aca cgc	528
Trp Pro Ala Asp Val Ala Glu Ala Met Ala Arg His Gly Xaa Thr Arg	
165 170 175	
ttg tat gcc gca ctg cac ctt ccc cct gag gtg ctg cta cca ccc ggc	576
Leu Tyr Ala Ala Leu His Leu Pro Pro Glu Val Leu Leu Pro Pro Gly	
180 185 190	
acc tac cac aca acc tcg tat ctc ctg att cac gac ggc gac cgc gct	624
Thr Tyr His Thr Thr Ser Tyr Leu Leu Ile His Asp Gly Asp Arg Ala	
195 200 205	
gtt gta act tac gag ggc gat act agt gcg ggc tat aat cat gat gtc	672
Val Val Thr Tyr Glu Gly Asp Thr Ser Ala Gly Tyr Asn His Asp Val	
210 215 220	
tcc ata ctt cgt gcg tgg atc cgt act aca aaa ata gtt ggt gat cat	720

Ser 225	Ile	Leu	Arg	Ala	Trp 230	Ile	Arg	Thr	Thr	Lys 235	Ile	Val	Gly	Asp	His 240	
ccg	ttg	gtc	ata	gag	cgt	gtg	cgg	gcc	att	gga	tgt	cat	ttt	gtg	ttg	768
Pro	Leu	Val	Ile	Glu	Arg	Val	Arg	Ala	Ile	Gly	Cys	His	Phe	Val	Leu	
				245					250					255		
ctg	ctc	acc	gca	gcc	cct	gag	cgg	tca	ccc	atg	cct	tat	gtt	cct	tac	816
Leu	Leu	Thr	Ala	Ala	Pro	Glu	Pro	Ser	Pro	Met	Pro	Tyr	Val	Pro	Tyr	
			260					265					270			
cct	cgt	tca	acg	gag	gtg	tat	gtc	cgg	tcc	ata	ttt	ggc	cct	ggc	ggc	864
Pro	Arg	Ser	Thr	Glu	Val	Tyr	Val	Arg	Ser	Ile	Phe	Gly	Pro	Gly	Gly	
			275				280					285				
tcc	cca	tcc	ttg	ttt	cgg	tca	gcc	tgc	tct	act	aaa	tct	act	ttc	cat	912
Ser	Pro	Ser	Leu	Phe	Pro	Ser	Ala	Cys	Ser	Thr	Lys	Ser	Thr	Phe	His	
	290						295				300					
gct	gtc	ccg	gtg	cat	atc	tgg	gat	cgg	ctc	atg	ctc	ttt	ggg	gcc	acc	960
Ala	Val	Pro	Val	His	Ile	Trp	Asp	Arg	Leu	Met	Leu	Phe	Gly	Ala	Thr	
305					310					315					320	
ctg	gac	gat	cag	gcg	ttt	tgc	tgt	tca	cgg	ctc	atg	act	tac	ctc	cgt	1008
Leu	Asp	Asp	Gln	Ala	Phe	Cys	Cys	Ser	Arg	Leu	Met	Thr	Tyr	Leu	Arg	
			325					330						335		
ggg	att	agt	tac	aag	gtc	act	gtc	ggc	gcg	ctt	gtc	gct	aat	gag	ggg	1056
Gly	Ile	Ser	Tyr	Lys	Val	Thr	Val	Gly	Ala	Leu	Val	Ala	Asn	Glu	Gly	
			340					345					350			
tgg	aac	gcc	tct	gaa	gac	gct	ctt	act	gca	rtg	atc	act	gca	gct	tat	1104
Trp	Asn	Ala	Ser	Glu	Asp	Ala	Leu	Thr	Ala	Xaa	Ile	Thr	Ala	Ala	Tyr	
	355						360					365				
ttg	act	att	tgc	cat	cag	cgt	tat	ctc	cgc	acc	cag	gcg	ata	tcc	aag	1152
Leu	Thr	Ile	Cys	His	Gln	Arg	Tyr	Leu	Arg	Thr	Gln	Ala	Ile	Ser	Lys	
	370					375					380					
ggc	atg	cgc	cgg	ttg	ggg	gtt	gag	cac	gcc	cag	aaa	ttt	atc	aca	aga	1200
Gly	Met	Arg	Arg	Leu	Gly	Val	Glu	His	Ala	Gln	Lys	Phe	Ile	Thr	Arg	
385				390						395					400	
ctc	tac	agt	tgg	cta	ttt	gag	aag	tct	ggc	cgt	gat	tat	atc	ccc	ggc	1248
Leu	Tyr	Ser	Trp	Leu	Phe	Glu	Lys	Ser	Gly	Arg	Asp	Tyr	Ile	Pro	Gly	
			405					410					415			
cgc	cag	ctt	cag	ttc	tat	gca	cag	tgc	cga	cgg	tgg	cta	tct	gca	ggc	1296
Arg	Gln	Leu	Gln	Phe	Tyr	Ala	Gln	Cys	Arg	Arg	Trp	Leu	Ser	Ala	Gly	
			420					425					430			
ttc	cac	cta	gac	ccc	agg	gta	ctt	gtt	ttt	gat	gag	tca	gta	cca	tgc	1344
Phe	His	Leu	Asp	Pro	Arg	Val	Leu	Val	Phe	Asp	Glu	Ser	Val	Pro	Cys	
		435					440					445				
cgc	tgt	agg	acg	ttt	ttg	aag	aaa	gtt	gcg	ggg	aaa	ttc	tgc	tgt	ttt	1392
Arg	Cys	Arg	Thr	Phe	Leu	Lys	Lys	Val	Ala	Gly	Lys	Phe	Cys	Cys	Phe	

450	455	460	
atg cgg tgg ctc ggg cag gag tgt acc tgc ttc ttg gag ccg gcc gag			1440
Met Arg Trp Leu Gly Gln Glu Cys Thr Cys Phe Leu Glu Pro Ala Glu			
465	470	475	480
ggt tta gtc ggc gat cat ggc cat gac aac gag gcc tat gag ggt tct			1488
Gly Leu Val Gly Asp His Gly His Asp Asn Glu Ala Tyr Glu Gly Ser			
	485	490	495
gag gtc gac ccg gct gaa cct gca cat ctt gat gtt tct ggg act tac			1536
Glu Val Asp Pro Ala Glu Pro Ala His Leu Asp Val Ser Gly Thr Tyr			
	500	505	510
gcc gtc cac ggg cac cag ctt gag gcc ctc tat agg gca ctt aat gtc			1584
Ala Val His Gly His Gln Leu Glu Ala Leu Tyr Arg Ala Leu Asn Val			
	515	520	525
cca caa gat att gcc gct cga gct tcc cga cta acg gca act gtt gag			1632
Pro Gln Asp Ile Ala Ala Arg Ala Ser Arg Leu Thr Ala Thr Val Glu			
	530	535	540
ctc gtt gca agt cca gac cgc tta gag tgc cgc acc gtg ctc ggt aat			1680
Leu Val Ala Ser Pro Asp Arg Leu Glu Cys Arg Thr Val Leu Gly Asn			
	545	550	555
aag acc ttc cgg acg acg gtg gtc gac ggc gcc cat cta gag gcg aat			1728
Lys Thr Phe Arg Thr Thr Val Val Asp Gly Ala His Leu Glu Ala Asn			
	565	570	575
ggc cct gag cag tat gtc tta tca ttt gac gcc tcc cgt cag tct atg			1776
Gly Pro Glu Gln Tyr Val Leu Ser Phe Asp Ala Ser Arg Gln Ser Met			
	580	585	590
ggg gcc ggg tgg cat agc ctc act tat gag ctc acc cct gct ggt ttg			1824
Gly Ala Gly Ser His Ser Leu Thr Tyr Glu Leu Thr Pro Ala Gly Leu			
	595	600	605
cag gtt agg att tca tct aat ggt ctg gat tgc act gct aca ttc ccc			1872
Gln Val Arg Ile Ser Ser Asn Gly Leu Asp Cys Thr Ala Thr Phe Pro			
	610	615	620
ccc ggt gga gcc cct agc gct gcg ccc ggg gag gtg gca gcc ttt tgc			1920
Pro Gly Gly Ala Pro Ser Ala Ala Pro Gly Glu Val Ala Ala Phe Cys			
	625	630	635
agt gcc ctt tat aga tat aac agg ttc acc cag cgg cac tgg ctg act			1968
Ser Ala Leu Tyr Arg Tyr Asn Arg Phe Thr Gln Arg His Ser Leu Thr			
	645	650	655
ggc gga tta tgg tta cac cct gag ggg ttg ctg ggt att ttc ccc cct			2016
Gly Gly Leu Trp Leu His Pro Glu Gly Leu Leu Gly Ile Phe Pro Pro			
	660	665	670
ttc tcc cct ggg cat atc tgg gag tct gcg aac ccc ttt tgc ggg gag			2064
Phe Ser Pro Gly His Ile Trp Glu Ser Ala Asn Pro Phe Cys Gly Glu			
	675	680	685

ggg act ttg tat acc cga act tgg tca aca tct ggc ttt tct agt gat	2112
Gly Thr Leu Tyr Thr Arg Thr Trp Ser Thr Ser Gly Phe Ser Ser Asp	
690 695 700	
ttc tcc ccc cct gaa gcg gcc gct cct gct atg gct gct acc ccg ggg	2160
Phe Ser Pro Pro Glu Ala Ala Pro Ala Met Ala Ala Thr Pro Gly	
705 710 715 720	
ctg ccc cat tct acc cca cct gtt agc gat att tgg gtg cta cca ccg	2208
Leu Pro His Ser Thr Pro Pro Val Ser Asp Ile Trp Val Leu Pro Pro	
725 730 735	
ccc tca gag gag ttt cag gtt gat gca gca cct gtg ccc cct gcc cct	2256
Pro Ser Glu Glu Phe Gln Val Asp Ala Ala Pro Val Pro Pro Ala Pro	
740 745 750	
gac cct gct gga ttg ccc ggt ccc gtt gtg ctt acc ccc ccc ccc cct	2304
Asp Pro Ala Gly Leu Pro Gly Pro Val Val Leu Thr Pro Pro Pro Pro	
755 760 765	
ccc cct gtg cat aag cca tca ata ccc ccg cct tcc cgt aac cgt cgt	2352
Pro Pro Val His Lys Pro Ser Ile Pro Pro Pro Ser Arg Asn Arg Arg	
770 775 780	
ctc ctc tat acc tat cct gac ggc gct aag gtg tat gca ggg tca ctg	2400
Leu Leu Tyr Thr Tyr Pro Asp Gly Ala Lys Val Tyr Ala Gly Ser Leu	
785 790 795 800	
ttt gaa tca gac tgt gac tgg ctg gtt aat gcc tca aac ccg ggc cat	2448
Phe Glu Ser Asp Cys Asp Trp Leu Val Asn Ala Ser Asn Pro Gly His	
805 810 815	
cgt ccc gga ggt ggc ctc tgc cat gcc ttt tac caa cgt ttt cca gaa	2496
Arg Pro Gly Gly Gly Leu Cys His Ala Phe Tyr Gln Arg Phe Pro Glu	
820 825 830	
gcg ttt tac cca act gaa ttc atc atg cgt gag ggt ctt gca gca tac	2544
Ala Phe Tyr Pro Thr Glu Phe Ile Met Arg Glu Gly Leu Ala Ala Tyr	
835 840 845	
acc ttg acc ccg cgc cct atc att cat gca gtc gct ccc gat tat agg	2592
Thr Leu Thr Pro Arg Pro Ile Ile His Ala Val Ala Pro Asp Tyr Arg	
850 855 860	
gtt gag cag aac ccg aag agg ctt gag gca gcg tac cgt gaa act tgt	2640
Val Glu Gln Asn Pro Lys Arg Leu Glu Ala Ala Tyr Arg Glu Thr Cys	
865 870 875 880	
tcc cgt cgt ggc acc gct gcc tac ccg ctt ttg ggt tgg ggt ata tac	2688
Ser Arg Arg Gly Thr Ala Ala Tyr Pro Leu Leu Gly Ser Gly Ile Tyr	
885 890 895	
cag gtc cct gtt agc ctc agt ttt gat gcc tgg gaa cgt aat cac cgc	2736
Gln Val Pro Val Ser Leu Ser Phe Asp Ala Trp Glu Arg Asn His Arg	
900 905 910	

ccc ggc gat gag ctt tac ttg acc gag ccc gct gca aat tgg ttt gag	2784
Pro Gly Asp Glu Leu Tyr Leu Thr Glu Pro Ala Ala Asn Trp Phe Glu	
915 920 925	
gct aat aag ccg gcg cag ccg gtg ctc acc ata act gag gac acg gcc	2832
Ala Asn Lys Pro Ala Gln Pro Val Leu Thr Ile Thr Glu Asp Thr Ala	
930 935 940	
cgt acg gcc aac ctg gca ttg gag att gat gcc gct aca gag gtc ggc	2880
Arg Thr Ala Asn Leu Ala Leu Glu Ile Asp Ala Ala Thr Glu Val Gly	
945 950 955 960	
cgt gct tgt gcc ggt tgc acc atc agc cct ggc att gtg cac tat cag	2928
Arg Ala Cys Ala Gly Cys Thr Ile Ser Pro Gly Ile Val His Tyr Gln	
965 970 975	
ttt acc gcc ggg gtc ccg ggc tcg ggc aag tca agg tcc ata caa cag	2976
Phe Thr Ala Gly Val Pro Gly Ser Gly Lys Ser Arg Ser Ile Gln Gln	
980 985 990	
gga gat gtc gat gtg gtg gtt gtg ccc acc cgg gag ctt cgt aat agt	3024
Gly Asp Val Asp Val Val Val Val Pro Thr Arg Glu Leu Arg Asn Ser	
995 1000 1005	
tgg cgc cgc cgg ggt ttt gcg gcc ttc aca ccc cac aca gcg gcc cgt	3072
Trp Arg Arg Arg Gly Phe Ala Ala Phe Thr Pro His Thr Ala Ala Arg	
1010 1015 1020	
gtt act atc ggc cgc cgc gtt gtg att gat gag gct cca tct ctc ccg	3120
Val Thr Ile Gly Arg Arg Val Val Ile Asp Glu Ala Pro Ser Leu Pro	
1025 1030 1035 1040	
cca cac ctg ttg ctg tta cat atg cag cgg gcc tcc tcg gtc cat ctc	3168
Pro His Leu Leu Leu His Met Gln Arg Ala Ser Ser Val His Leu	
1045 1050 1055	
ctc ggt gac cca aat cag atc cct gct att gat ttt gag cac gcc ggc	3216
Leu Gly Asp Pro Asn Gln Ile Pro Ala Ile Asp Phe Glu His Ala Gly	
1060 1065 1070	
ctg gtc cct gcg atc cgt ccc gag ctt gcg cca acg agc tgg tgg crc	3264
Leu Val Pro Ala Ile Arg Pro Glu Leu Ala Pro Thr Ser Trp Trp Xaa	
1075 1080 1085	
gtt aca cac cgt tgc ccg gcc gat gtg tgc gag ctc ata cgc gga gcc	3312
Val Thr His Arg Cys Pro Ala Asp Val Cys Glu Leu Ile Arg Gly Ala	
1090 1095 1100	
tac cct aaa atc cag acc acg agc cgt gtg cta cgg tcc ctg ttt tgg	3360
Tyr Pro Lys Ile Gln Thr Thr Ser Arg Val Leu Arg Ser Leu Phe Trp	
1105 1110 1115 1120	
aat gaa ccg gcc att ggc cag aag ttg gtt ytc acg cag gcg gca aag	3408
Asn Glu Pro Ala Ile Gly Gln Lys Leu Val Xaa Thr Gln Ala Ala Lys	
1125 1130 1135	
gct gct aac cct ggt gcg att acg gtc cac gaa gct cag ggt gcc acc	3456

Ala	Ala	Asn	Pro	Gly	Ala	Ile	Thr	Val	His	Glu	Ala	Gln	Gly	Ala	Thr		
			1140					1145					1150				
ttc	aca	gag	acc	aca	atc	ata	gcc	acg	gcc	gac	gcc	agg	ggc	ctt	atc	3504	
Phe	Thr	Glu	Thr	Thr	Ile	Ile	Ala	Thr	Ala	Asp	Ala	Arg	Gly	Leu	Ile		
		1155					1160					1165					
cag	tca	tcc	cgg	gct	cat	gct	ata	gtt	gca	ctt	act	cgc	cac	act	gag	3552	
Gln	Ser	Ser	Arg	Ala	His	Ala	Ile	Val	Ala	Leu	Thr	Arg	His	Thr	Glu		
		1170				1175					1180						
aag	tgt	gtt	atc	ctg	gat	gcc	ccc	ggc	ctg	ctt	cgt	gag	gtc	ggc	att	3600	
Lys	Cys	Val	Ile	Leu	Asp	Ala	Pro	Gly	Leu	Leu	Arg	Glu	Val	Gly	Ile		
1185					1190				1195					1200			
tcg	gat	gtg	att	gtc	aac	aac	ttt	ttc	ctt	gct	ggc	gag	gtc	ggc		3648	
Ser	Asp	Val	Ile	Val	Asn	Asn	Phe	Phe	Leu	Ala	Gly	Gly	Glu	Val	Gly		
				1205				1210					1215				
crc	cac	cgc	cct	tct	gtg	ata	cct	cgc	ggc	aac	cct	gat	caa	aac	ctc	3696	
Xaa	His	Arg	Pro	Ser	Val	Ile	Pro	Arg	Gly	Asn	Pro	Asp	Gln	Asn	Leu		
			1220				1225						1230				
ggg	act	tta	cag	gcc	ttc	ccg	ccg	tcc	tgt	caa	att	agt	gct	tac	cat	3744	
Gly	Thr	Leu	Gln	Ala	Phe	Pro	Pro	Ser	Cys	Gln	Ile	Ser	Ala	Tyr	His		
		1235				1240					1245						
cag	ttg	gct	gag	gaa	ctg	ggc	cat	cgc	ccg	gcc	cct	gtc	gcc	gcc	gtc	3792	
Gln	Leu	Ala	Glu	Glu	Leu	Gly	His	Arg	Pro	Ala	Pro	Val	Ala	Ala	Val		
		1250				1255					1260						
ttg	ccc	cct	tgc	cct	gag	ctt	gag	cag	ggc	ctg	ctc	tac	atg	cca	cag	3840	
Leu	Pro	Pro	Cys	Pro	Glu	Leu	Glu	Gln	Gly	Leu	Leu	Tyr	Met	Pro	Gln		
1265					1270				1275				1280				
gag	ctc	act	gtg	tcc	gat	agt	gtg	ttg	gtt	ttt	gag	ctt	acg	gat	ata	3888	
Glu	Leu	Thr	Val	Ser	Asp	Ser	Val	Leu	Val	Phe	Glu	Leu	Thr	Asp	Ile		
				1285				1290					1295				
gtt	cat	tgc	cgc	atg	gcc	gct	cca	agc	cag	cga	aag	gct	gtt	ctc	tca	3936	
Val	His	Cys	Arg	Met	Ala	Ala	Pro	Ser	Gln	Arg	Lys	Ala	Val	Leu	Ser		
			1300				1305						1310				
aca	ctt	gtg	ggg	agg	tat	ggc	cgt	agg	acg	aaa	cta	tat	gag	gcg	gcg	3984	
Thr	Leu	Val	Gly	Arg	Tyr	Gly	Arg	Arg	Thr	Lys	Leu	Tyr	Glu	Ala	Ala		
		1315				1320						1325					
cat	tca	gat	gtt	cgt	gag	tcc	cta	gct	agg	ttc	atc	cct	act	atc	ggg	4032	
His	Ser	Asp	Val	Arg	Glu	Ser	Leu	Ala	Arg	Phe	Ile	Pro	Thr	Ile	Gly		
		1330				1335					1340						
cct	gtt	cag	gct	acc	aca	tgt	gag	ttg	tat	gag	ttg	gtt	gag	gct	atg	4080	
Pro	Val	Gln	Ala	Thr	Thr	Cys	Glu	Leu	Tyr	Glu	Leu	Val	Glu	Ala	Met		
1345					1350					1355				1360			
gtg	gag	aaa	ggc	cag	gac	ggc	tct	gca	gtc	tta	gag	ctt	gat	ctt	tgt	4128	
Val	Glu	Lys	Gly	Gln	Asp	Gly	Ser	Ala	Val	Leu	Glu	Leu	Asp	Leu	Cys		

1365	1370	1375	
aat cgt gat gtc tcg cgc atc aca ttt ttc caa aaa gwc tgc aac aag			4176
Asn Arg Asp Val Ser Arg Ile Thr Phe Phe Gln Lys Xaa Cys Asn Lys			
1380	1385	1390	
ttt aca act ggt gag acc atc gcc cac ggc aag gtt ggc cag ggt ata			4224
Phe Thr Thr Gly Glu Thr Ile Ala His Gly Lys Val Gly Gln Gly Ile			
1395	1400	1405	
tcg gcc tgg agt aag acc ttc tgc gct ctg ttc ggc cgc tgg ttc cgc			4272
Ser Ala Trp Ser Lys Thr Phe Cys Ala Leu Phe Gly Pro Trp Phe Arg			
1410	1415	1420	
gcc att gaa aaa gaa ata ttg gcc ctg ctc ccg cct aat atc ttt tat			4320
Ala Ile Glu Lys Glu Ile Leu Ala Leu Leu Pro Pro Asn Ile Phe Tyr			
1425	1430	1435	1440
ggc gac gct tat gag gag tca gtt ttt gcc gcc gct gtg tcc ggg gcg			4368
Gly Asp Ala Tyr Glu Glu Ser Val Phe Ala Ala Val Ser Gly Ala			
1445	1450	1455	
ggg tca tgt atg gta ttt gaa aat gac ttt tca gag ttt gac agt acc			4416
Gly Ser Cys Met Val Phe Glu Asn Asp Phe Ser Glu Phe Asp Ser Thr			
1460	1465	1470	
cag aat aat ttc tct ctt ggc ctt gag tgt gtg gtt atg gag gag tgc			4464
Gln Asn Asn Phe Ser Leu Gly Leu Glu Cys Val Val Met Glu Glu Cys			
1475	1480	1485	
ggc atg cct caa tgg cta att agg ttg tac cat ctg gtt cgg tct gcc			4512
Gly Met Pro Gln Trp Leu Ile Arg Leu Tyr His Leu Val Arg Ser Ala			
1490	1495	1500	
tgg att ctg cag gcg ccg aag gag tct ctt aag ggt ttc tgg aag aag			4560
Trp Ile Leu Gln Ala Pro Lys Glu Ser Leu Lys Gly Phe Trp Lys Lys			
1505	1510	1515	1520
cat tct ggt gag cct ggt acc ctt ctt tgg aat acc gtc tgg aat atg			4608
His Ser Gly Glu Pro Gly Thr Leu Leu Trp Asn Thr Val Trp Asn Met			
1525	1530	1535	
gcg att ata gca cat tgc tat gag ttc cgt gac ttt cgt gtt gct gcc			4656
Ala Ile Ile Ala His Cys Tyr Glu Phe Arg Asp Phe Arg Val Ala Ala			
1540	1545	1550	
ttt aag ggt gat gat tcg gtg gtc ctc tgt agt gac tac cga cag agc			4704
Phe Lys Gly Asp Asp Ser Val Val Leu Cys Ser Asp Tyr Arg Gln Ser			
1555	1560	1565	
cgc aat gca gct gcc tta att gct ggc tgt ggg ctc aaa ttg aag gtt			4752
Arg Asn Ala Ala Ala Leu Ile Ala Gly Cys Gly Leu Lys Leu Lys Val			
1570	1575	1580	
gat tac cgc cct atc ggg ctg tat gct ggg gtg gtg gtg gcc ccc ggt			4800
Asp Tyr Arg Pro Ile Gly Leu Tyr Ala Gly Val Val Val Ala Pro Gly			
1585	1590	1595	1600

ttg ggg aca ctg ccc gat gtg gtg cgt ttt gct ggt cgg ttg tct gaa	4848
Leu Gly Thr Leu Pro Asp Val Val Arg Phe Ala Gly Arg Leu Ser Glu	
1605 1610 1615	
aag aat tgg ggc ccc ggc ccg gaa cgt gct gag cag ctg cgt ctt gct	4896
Lys Asn Trp Gly Pro Gly Pro Glu Arg Ala Glu Gln Leu Arg Leu Ala	
1620 1625 1630	
gtc tgc gac ttc ctt cga ggg ttg acg aat gtt gcg cag gtc tgt gtt	4944
Val Cys Asp Phe Leu Arg Gly Leu Thr Asn Val Ala Gln Val Cys Val	
1635 1640 1645	
gat gtt gtg tcc cgt gtc tat gga gtc agc ccc ggg ctc gta cat aac	4992
Asp Val Val Ser Arg Val Tyr Gly Val Ser Pro Gly Leu Val His Asn	
1650 1655 1660	
ctt att ggc atg ctg cag acc atc gcc gat ggc aag gcc cac ttt aca	5040
Leu Ile Gly Met Leu Gln Thr Ile Ala Asp Gly Lys Ala His Phe Thr	
1665 1670 1675 1680	
gag act att aaa cct gta ctt gat ctc aca aat tcc atc ata cag cgg	5088
Glu Thr Ile Lys Pro Val Leu Asp Leu Thr Asn Ser Ile Ile Gln Arg	
1685 1690 1695	
gtg gaa tgaataacat gtcttttgca tcgcccattgg gatcacc atg cgc cct agg	5143
Val Glu Met Arg Pro Arg	
1700	
gct gtt ctg ttg ttg ttc ctc atg ttt ctg cct atg ctg ccc gcg cca	5191
Ala Val Leu Leu Leu Phe Leu Met Phe Leu Pro Met Leu Pro Ala Pro	
1705 1710 1715	
ccg gcc ggt cag ccg tct ggc cgt cgc cgt ggg cgg cgc agc ggc ggt	5239
Pro Ala Gly Gln Pro Ser Gly Arg Arg Gly Arg Arg Ser Gly Gly	
1720 1725 1730	
gcc ggc ggt ggt ttc tgg agt gac agg gtt gat tct cag ccc ttc gcc	5287
Ala Gly Gly Gly Phe Trp Ser Asp Arg Val Asp Ser Gln Pro Phe Ala	
1735 1740 1745 1750	
ctc ccc tat att cat cca acc aac ccc ttc gcc gcc gat gtc gtt toa	5335
Leu Pro Tyr Ile His Pro Thr Asn Pro Phe Ala Ala Asp Val Val Ser	
1755 1760 1765	
caa ccc ggg gct gga act cgc cct cga cag ccg ccc cgc ccc ctc ggt	5383
Gln Pro Gly Ala Gly Thr Arg Pro Arg Gln Pro Pro Arg Pro Leu Gly	
1770 1775 1780	
tcc gct tgg cgt gac cag tcc aag cgc ccc tcc gtt gcc ccc cgt cgt	5431
Ser Ala Trp Arg Asp Gln Ser Lys Arg Pro Ser Val Ala Pro Arg Arg	
1785 1790 1795	
cga tct acc cca gct ggg gct gcg ccg cta act gcc ata tca cca gcc	5479
Arg Ser Thr Pro Ala Gly Ala Ala Pro Leu Thr Ala Ile Ser Pro Ala	
1800 1805 1810	

cct gat aca gct cct gta cct gat gtt gac tca cgt ggt gct att ttg	5527
Pro Asp Thr Ala Pro Val Pro Asp Val Asp Ser Arg Gly Ala Ile Leu	
1815 1820 1825 1830	
cgc cgg cag tac aat ttg tct acg tcc cgg ctt aca tca tct gtt gct	5575
Arg Arg Gln Tyr Asn Leu Ser Thr Ser Pro Leu Thr Ser Ser Val Ala	
1835 1840 1845	
tct ggt act aat ctg gtt ctc tat gct gcc cgg ctg aac cct ctc ttg	5623
Ser Gly Thr Asn Leu Val Leu Tyr Ala Ala Pro Leu Asn Pro Leu Leu	
1850 1855 1860	
cct ctt cag gat ggc acc aac act cat att atg gct act gag gca tct	5671
Pro Leu Gln Asp Gly Thr Asn Thr His Ile Met Ala Thr Glu Ala Ser	
1865 1870 1875	
aat tac gcc cag tat cgg gtt gtt cgg gct acg att cgt tat cgc cgg	5719
Asn Tyr Ala Gln Tyr Arg Val Val Arg Ala Thr Ile Arg Tyr Arg Pro	
1880 1885 1890	
ttg gtg cca aat gct gtt ggt ggt tat gct atc tct att tct ttc tgg	5767
Leu Val Pro Asn Ala Val Gly Gly Tyr Ala Ile Ser Ile Ser Phe Trp	
1895 1900 1905 1910	
cct caa act aca act acc cct act tct gtt gac atg aat tct atc act	5815
Pro Gln Thr Thr Thr Thr Pro Thr Ser Val Asp Met Asn Ser Ile Thr	
1915 1920 1925	
tct act gat gtc agg atc ttg gtc cag ccc ggt ata gcc tcc gag tta	5863
Ser Thr Asp Val Arg Ile Leu Val Gln Pro Gly Ile Ala Ser Glu Leu	
1930 1935 1940	
gtc atc cct agt gaa cgc ctt cac tac cgc aac caa ggc tgg cgc tct	5911
Val Ile Pro Ser Glu Arg Leu His Tyr Arg Asn Gln Gly Trp Arg Ser	
1945 1950 1955	
gtt gag acc acg ggt gtg gcc gaa gag gag gct acc tcc ggt ctg gta	5959
Val Glu Thr Thr Gly Val Ala Glu Glu Glu Ala Thr Ser Gly Leu Val	
1960 1965 1970	
atg ctt tgt att cat ggc tcc cct gtt aac tcc tac act aat aca cct	6007
Met Leu Cys Ile His Gly Ser Pro Val Asn Ser Tyr Thr Asn Thr Pro	
1975 1980 1985 1990	
tac acc ggt gca ttg ggg ctt ctt gat ttt gca tta gaa ctt gaa ttt	6055
Tyr Thr Gly Ala Leu Gly Leu Leu Asp Phe Ala Leu Glu Leu Glu Phe	
1995 2000 2005	
aga aat ttg aca ccc ggg aac act aac acc cgt gtt tcc cgg tat act	6103
Arg Asn Leu Thr Pro Gly Asn Thr Asn Thr Arg Val Ser Arg Tyr Thr	
2010 2015 2020	
agc aca gcc cgc cac cgg ctg cgc cgc ggt gct gat ggg acc gct gag	6151
Ser Thr Ala Arg His Arg Leu Arg Arg Gly Ala Asp Gly Thr Ala Glu	
2025 2030 2035	
ctc acc acc aca gca gcc aca cgc ttc atg aag gat ttg cat ttt act	6199

Leu	Thr	Thr	Thr	Ala	Ala	Thr	Arg	Phe	Met	Lys	Asp	Leu	His	Phe	Thr		
2040						2045					2050						
ggc	acg	aac	ggc	gtt	ggc	gag	gtg	ggc	cgt	ggc	att	gcc	ctg	act	ctg	6247	
Gly	Thr	Asn	Gly	Val	Gly	Glu	Val	Gly	Arg	Gly	Ile	Ala	Leu	Thr	Leu		
2055					2060					2065					2070		
ttt	aat	ctt	gct	gat	acg	ctt	ctt	ggc	ggc	tta	cgc	aca	gaa	ttg	att	6295	
Phe	Asn	Leu	Ala	Asp	Thr	Leu	Leu	Gly	Gly	Leu	Pro	Thr	Glu	Leu	Ile		
				2075					2080					2085			
tcg	tcg	gct	ggg	ggc	caa	ctg	ttt	tac	tcc	cgc	cct	gtt	gtc	tcg	gcc	6343	
Ser	Ser	Ala	Gly	Gly	Gln	Leu	Phe	Tyr	Ser	Arg	Pro	Val	Val	Ser	Ala		
			2090					2095					2100				
aat	ggc	gag	cca	aca	gta	aag	tta	tac	aca	tct	gtt	gag	aat	gcg	cag	6391	
Asn	Gly	Glu	Pro	Thr	Val	Lys	Leu	Tyr	Thr	Ser	Val	Glu	Asn	Ala	Gln		
		2105					2110					2115					
caa	gac	aag	ggc	atc	acc	att	cca	cac	gac	ata	gat	tta	ggc	gac	tcc	6439	
Gln	Asp	Lys	Gly	Ile	Thr	Ile	Pro	His	Asp	Ile	Asp	Leu	Gly	Asp	Ser		
		2120				2125						2130					
cgt	gtg	gtt	atc	cag	gat	tat	gat	aac	cag	cac	gaa	caa	gat	cga	cct	6487	
Arg	Val	Val	Ile	Gln	Asp	Tyr	Asp	Asn	Gln	His	Glu	Gln	Asp	Arg	Pro		
				2140					2145						2150		
acc	cgc	tca	cct	gcc	ccc	tcc	cgc	cct	tcc	tca	gtt	ctt	cgt	gcc	aat	6535	
Thr	Pro	Ser	Pro	Ala	Pro	Ser	Arg	Pro	Phe	Ser	Val	Leu	Arg	Ala	Asn		
				2155					2160					2165			
gat	gtt	ttg	tgg	ctc	tct	ctc	act	gcc	gct	gag	tac	gac	cag	acc	acg	6583	
Asp	Val	Leu	Trp	Leu	Ser	Leu	Thr	Ala	Ala	Glu	Tyr	Xaa	Gln	Thr	Thr		
			2170					2175					2180				
tat	ggg	tcg	tcc	acc	aac	cct	atg	tat	gtc	tct	gat	aca	gtc	acg	ctt	6631	
Tyr	Gly	Ser	Ser	Thr	Asn	Pro	Met	Tyr	Val	Ser	Asp	Thr	Val	Thr	Leu		
		2185					2190					2195					
gtt	aat	gta	gcc	act	ggc	gct	cag	gct	gtt	gcc	cgc	tct	ctt	gac	tgg	6679	
Val	Asn	Val	Ala	Thr	Gly	Ala	Gln	Ala	Val	Ala	Arg	Ser	Leu	Asp	Trp		
		2200				2205					2210						
tct	aaa	gtt	act	ctg	gat	ggc	cgc	cct	ctt	act	acc	att	cag	cag	tat	6727	
Ser	Lys	Val	Thr	Leu	Asp	Gly	Arg	Pro	Leu	Thr	Thr	Ile	Gln	Gln	Tyr		
				2215		2220				2225				2230			
tct	aag	aaa	ttt	tat	gtt	ctc	cgc	ctt	cgs	ggg	aag	ctg	tcc	ttt	tgg	6775	
Ser	Lys	Lys	Phe	Tyr	Val	Leu	Pro	Leu	Xaa	Gly	Lys	Leu	Ser	Phe	Trp		
			2235					2240					2245				
gag	gct	ggc	acg	acc	aag	gcc	ggc	tac	cgc	tat	aat	tat	aat	acc	act	6823	
Glu	Ala	Gly	Thr	Thr	Lys	Ala	Gly	Tyr	Pro	Tyr	Asn	Tyr	Asn	Thr	Thr		
			2250				2255						2260				
gct	agt	gac	caa	att	ttg	att	gag	aac	gcg	gcc	ggc	cac	cgt	gtc	gcc	6871	
Ala	Ser	Asp	Gln	Ile	Leu	Ile	Glu	Asn	Ala	Ala	Gly	His	Arg	Val	Ala		

2265	2270	2275	
att tct act tat acc act agt ttg ggt gcc ggc cct acc tcg aty tct			6919
Ile Ser Thr Tyr Thr Thr Ser Leu Gly Ala Gly Pro Thr Ser Xaa Ser			
2280	2285	2290	
gcg gtc ggt gta cta gct cca cat tcg gcc ctt gct gtt ctc gag gat			6967
Ala Val Gly Val Leu Ala Pro His Ser Ala Leu Ala Val Leu Glu Asp			
2295	2300	2305	2310
act gtt gat tat cct gct cgt gcc cat act ttt gat gat ttc tgc ccg			7015
Thr Val Asp Tyr Pro Ala Arg Ala His Thr Phe Asp Asp Phe Cys Pro			
	2315	2320	2325
gag tgt cgc acc ctt ggt ctg cag ggt tgt gca ttc caa tct act att			7063
Glu Cys Arg Thr Leu Gly Leu Gln Gly Cys Ala Phe Gln Ser Thr Ile			
	2330	2335	2340
gct gaa ctt cag cgt ctt aaa atg aag gta ggt aaa acc cgg gag tct			7111
Ala Glu Leu Thr Arg Leu Lys Met Lys Val Gly Lys Thr Arg Glu Ser			
	2345	2350	2355
taattaattc cttttgtgcc cccttcgcag ttctctttgg ctttatttct catttctgct			7171
ttccgcgctc cctggaaaaa aaaaaaaaaa a			7202

<210> 91
 <211> 1698
 <212> PRT
 <213> Hepatitis E Virus

<220>
 <223> Xaa = Unknown or Other at position 174
 <223> Xaa = Unknown or Other at position 363
 <223> Xaa = Unknown or Other at position 1088
 <223> Xaa = Unknown or Other at position 1131
 <223> Xaa = Unknown or Other at position 1217
 <223> Xaa = Unknown or Other at position 1389

<400> 91
 Pro Gly Ile Thr Thr Ala Ile Glu Gln Ala Ala Leu Ala Ala Ala Asn
 1 5 10 15
 Ser Ala Leu Ala Asn Ala Val Val Val Arg Pro Phe Leu Ser Arg Val
 20 25 30
 Gln Thr Glu Ile Leu Ile Asn Leu Met Gln Pro Arg Gln Leu Val Phe
 35 40 45
 Arg Pro Glu Val Leu Trp Asn His Pro Ile Gln Arg Val Ile His Asn
 50 55 60
 Glu Leu Glu Gln Tyr Cys Arg Ala Arg Ala Gly Arg Cys Leu Glu Val
 65 70 75 80
 Gly Ala His Pro Arg Ser Ile Asn Asp Asn Pro Asn Val Leu His Arg
 85 90 95
 Cys Phe Leu Arg Pro Val Gly Arg Asp Val Gln Arg Trp Tyr Ser Ala

			100					105				110			
Pro	Thr	Arg	Gly	Pro	Ala	Ala	Asn	Cys	Arg	Arg	Ser	Ala	Leu	Arg	Gly
		115					120					125			
Leu	Pro	Pro	Ala	Asp	Arg	Thr	Tyr	Cys	Phe	Asp	Gly	Phe	Ser	Arg	Cys
	130					135					140				
Ala	Phe	Ala	Ala	Glu	Thr	Gly	Val	Ala	Leu	Tyr	Ser	Leu	His	Asp	Leu
145				150						155					160
Trp	Pro	Ala	Asp	Val	Ala	Glu	Ala	Met	Ala	Arg	His	Gly	Xaa	Thr	Arg
			165					170						175	
Leu	Tyr	Ala	Ala	Leu	His	Leu	Pro	Pro	Glu	Val	Leu	Leu	Pro	Pro	Gly
		180					185						190		
Thr	Tyr	His	Thr	Thr	Ser	Tyr	Leu	Leu	Ile	His	Asp	Gly	Asp	Arg	Ala
		195					200					205			
Val	Val	Thr	Tyr	Glu	Gly	Asp	Thr	Ser	Ala	Gly	Tyr	Asn	His	Asp	Val
	210					215					220				
Ser	Ile	Leu	Arg	Ala	Trp	Ile	Arg	Thr	Thr	Lys	Ile	Val	Gly	Asp	His
225				230						235					240
Pro	Leu	Val	Ile	Glu	Arg	Val	Arg	Ala	Ile	Gly	Cys	His	Phe	Val	Leu
			245					250						255	
Leu	Leu	Thr	Ala	Ala	Pro	Glu	Pro	Ser	Pro	Met	Pro	Tyr	Val	Pro	Tyr
		260					265						270		
Pro	Arg	Ser	Thr	Glu	Val	Tyr	Val	Arg	Ser	Ile	Phe	Gly	Pro	Gly	Gly
		275					280					285			
Ser	Pro	Ser	Leu	Phe	Pro	Ser	Ala	Cys	Ser	Thr	Lys	Ser	Thr	Phe	His
	290					295					300				
Ala	Val	Pro	Val	His	Ile	Trp	Asp	Arg	Leu	Met	Leu	Phe	Gly	Ala	Thr
305				310						315					320
Leu	Asp	Asp	Gln	Ala	Phe	Cys	Cys	Ser	Arg	Leu	Met	Thr	Tyr	Leu	Arg
			325					330						335	
Gly	Ile	Ser	Tyr	Lys	Val	Thr	Val	Gly	Ala	Leu	Val	Ala	Asn	Glu	Gly
		340					345						350		
Trp	Asn	Ala	Ser	Glu	Asp	Ala	Leu	Thr	Ala	Xaa	Ile	Thr	Ala	Ala	Tyr
	355					360						365			
Leu	Thr	Ile	Cys	His	Gln	Arg	Tyr	Leu	Arg	Thr	Gln	Ala	Ile	Ser	Lys
	370					375					380				
Gly	Met	Arg	Arg	Leu	Gly	Val	Glu	His	Ala	Gln	Lys	Phe	Ile	Thr	Arg
385				390						395					400
Leu	Tyr	Ser	Trp	Leu	Phe	Glu	Lys	Ser	Gly	Arg	Asp	Tyr	Ile	Pro	Gly
			405					410					415		
Arg	Gln	Leu	Gln	Phe	Tyr	Ala	Gln	Cys	Arg	Arg	Trp	Leu	Ser	Ala	Gly
		420					425					430			
Phe	His	Leu	Asp	Pro	Arg	Val	Leu	Val	Phe	Asp	Glu	Ser	Val	Pro	Cys
	435						440					445			
Arg	Cys	Arg	Thr	Phe	Leu	Lys	Lys	Val	Ala	Gly	Lys	Phe	Cys	Cys	Phe
	450				455					460					
Met	Arg	Trp	Leu	Gly	Gln	Glu	Cys	Thr	Cys	Phe	Leu	Glu	Pro	Ala	Glu
465				470						475					480
Gly	Leu	Val	Gly	Asp	His	Gly	His	Asp	Asn	Glu	Ala	Tyr	Glu	Gly	Ser
			485					490					495		
Glu	Val	Asp	Pro	Ala	Glu	Pro	Ala	His	Leu	Asp	Val	Ser	Gly	Thr	Tyr
		500					505						510		
Ala	Val	His	Gly	His	Gln	Leu	Glu	Ala	Leu	Tyr	Arg	Ala	Leu	Asn	Val
	515					520						525			
Pro	Gln	Asp	Ile	Ala	Ala	Arg	Ala	Ser	Arg	Leu	Thr	Ala	Thr	Val	Glu
	530				535						540				
Leu	Val	Ala	Ser	Pro	Asp	Arg	Leu	Glu	Cys	Arg	Thr	Val	Leu	Gly	Asn
545					550					555					560

Lys	Thr	Phe	Arg	Thr	Thr	Val	Val	Asp	Gly	Ala	His	Leu	Glu	Ala	Asn
				565					570					575	
Gly	Pro	Glu	Gln	Tyr	Val	Leu	Ser	Phe	Asp	Ala	Ser	Arg	Gln	Ser	Met
			580					585					590		
Gly	Ala	Gly	Ser	His	Ser	Leu	Thr	Tyr	Glu	Leu	Thr	Pro	Ala	Gly	Leu
		595					600					605			
Gln	Val	Arg	Ile	Ser	Ser	Asn	Gly	Leu	Asp	Cys	Thr	Ala	Thr	Phe	Pro
	610					615					620				
Pro	Gly	Gly	Ala	Pro	Ser	Ala	Ala	Pro	Gly	Glu	Val	Ala	Ala	Phe	Cys
	625				630					635					640
Ser	Ala	Leu	Tyr	Arg	Tyr	Asn	Arg	Phe	Thr	Gln	Arg	His	Ser	Leu	Thr
			645					650						655	
Gly	Gly	Leu	Trp	Leu	His	Pro	Glu	Gly	Leu	Leu	Gly	Ile	Phe	Pro	Pro
		660					665						670		
Phe	Ser	Pro	Gly	His	Ile	Trp	Glu	Ser	Ala	Asn	Pro	Phe	Cys	Gly	Glu
		675					680					685			
Gly	Thr	Leu	Tyr	Thr	Arg	Thr	Trp	Ser	Thr	Ser	Gly	Phe	Ser	Ser	Asp
	690					695					700				
Phe	Ser	Pro	Pro	Glu	Ala	Ala	Ala	Pro	Ala	Met	Ala	Ala	Thr	Pro	Gly
	705				710					715					720
Leu	Pro	His	Ser	Thr	Pro	Pro	Val	Ser	Asp	Ile	Trp	Val	Leu	Pro	Pro
			725						730					735	
Pro	Ser	Glu	Glu	Phe	Gln	Val	Asp	Ala	Ala	Pro	Val	Pro	Pro	Ala	Pro
		740					745							750	
Asp	Pro	Ala	Gly	Leu	Pro	Gly	Pro	Val	Val	Leu	Thr	Pro	Pro	Pro	Pro
		755				760						765			
Pro	Pro	Val	His	Lys	Pro	Ser	Ile	Pro	Pro	Pro	Ser	Arg	Asn	Arg	Arg
	770					775					780				
Leu	Leu	Tyr	Thr	Tyr	Pro	Asp	Gly	Ala	Lys	Val	Tyr	Ala	Gly	Ser	Leu
	785				790					795					800
Phe	Glu	Ser	Asp	Cys	Asp	Trp	Leu	Val	Asn	Ala	Ser	Asn	Pro	Gly	His
			805						810					815	
Arg	Pro	Gly	Gly	Gly	Leu	Cys	His	Ala	Phe	Tyr	Gln	Arg	Phe	Pro	Glu
			820					825						830	
Ala	Phe	Tyr	Pro	Thr	Glu	Phe	Ile	Met	Arg	Glu	Gly	Leu	Ala	Ala	Tyr
		835					840					845			
Thr	Leu	Thr	Pro	Arg	Pro	Ile	Ile	His	Ala	Val	Ala	Pro	Asp	Tyr	Arg
	850					855					860				
Val	Glu	Gln	Asn	Pro	Lys	Arg	Leu	Glu	Ala	Ala	Tyr	Arg	Glu	Thr	Cys
	865				870					875					880
Ser	Arg	Arg	Gly	Thr	Ala	Ala	Tyr	Pro	Leu	Leu	Gly	Ser	Gly	Ile	Tyr
			885					890						895	
Gln	Val	Pro	Val	Ser	Leu	Ser	Phe	Asp	Ala	Trp	Glu	Arg	Asn	His	Arg
		900						905					910		
Pro	Gly	Asp	Glu	Leu	Tyr	Leu	Thr	Glu	Pro	Ala	Ala	Asn	Trp	Phe	Glu
		915					920					925			
Ala	Asn	Lys	Pro	Ala	Gln	Pro	Val	Leu	Thr	Ile	Thr	Glu	Asp	Thr	Ala
	930					935					940				
Arg	Thr	Ala	Asn	Leu	Ala	Leu	Glu	Ile	Asp	Ala	Ala	Thr	Glu	Val	Gly
	945				950					955					960
Arg	Ala	Cys	Ala	Gly	Cys	Thr	Ile	Ser	Pro	Gly	Ile	Val	His	Tyr	Gln
			965						970					975	
Phe	Thr	Ala	Gly	Val	Pro	Gly	Ser	Gly	Lys	Ser	Arg	Ser	Ile	Gln	Gln
		980					985						990		
Gly	Asp	Val	Asp	Val	Val	Val	Val	Pro	Thr	Arg	Glu	Leu	Arg	Asn	Ser
		995					1000					1005			
Trp	Arg	Arg	Arg	Gly	Phe	Ala	Ala	Phe	Thr	Pro	His	Thr	Ala	Ala	Arg

1010	1015	1020
Val Thr Ile Gly Arg Arg Val Val Ile Asp Glu Ala Pro Ser Leu Pro		
1025	1030	1035
Pro His Leu Leu Leu Leu His Met Gln Arg Ala Ser Ser Val His Leu		1040
	1045	1050
Leu Gly Asp Pro Asn Gln Ile Pro Ala Ile Asp Phe Glu His Ala Gly		1055
	1060	1065
Leu Val Pro Ala Ile Arg Pro Glu Leu Ala Pro Thr Ser Trp Trp Xaa		1070
	1075	1080
Val Thr His Arg Cys Pro Ala Asp Val Cys Glu Leu Ile Arg Gly Ala		1085
	1090	1095
Tyr Pro Lys Ile Gln Thr Thr Ser Arg Val Leu Arg Ser Leu Phe Trp		1100
1105	1110	1115
Asn Glu Pro Ala Ile Gly Gln Lys Leu Val Xaa Thr Gln Ala Ala Lys		1120
	1125	1130
Ala Ala Asn Pro Gly Ala Ile Thr Val His Glu Ala Gln Gly Ala Thr		1135
	1140	1145
Phe Thr Glu Thr Thr Ile Ile Ala Thr Ala Asp Ala Arg Gly Leu Ile		1150
	1155	1160
Gln Ser Ser Arg Ala His Ala Ile Val Ala Leu Thr Arg His Thr Glu		1165
	1170	1175
Lys Cys Val Ile Leu Asp Ala Pro Gly Leu Leu Arg Glu Val Gly Ile		1180
1185	1190	1195
Ser Asp Val Ile Val Asn Asn Phe Phe Leu Ala Gly Gly Glu Val Gly		1200
	1205	1210
Xaa His Arg Pro Ser Val Ile Pro Arg Gly Asn Pro Asp Gln Asn Leu		1215
	1220	1225
Gly Thr Leu Gln Ala Phe Pro Pro Ser Cys Gln Ile Ser Ala Tyr His		1230
	1235	1240
Gln Leu Ala Glu Glu Leu Gly His Arg Pro Ala Pro Val Ala Ala Val		1245
	1250	1255
Leu Pro Pro Cys Pro Glu Leu Glu Gln Gly Leu Leu Tyr Met Pro Gln		1260
1265	1270	1275
Glu Leu Thr Val Ser Asp Ser Val Leu Val Phe Glu Leu Thr Asp Ile		1280
	1285	1290
Val His Cys Arg Met Ala Ala Pro Ser Gln Arg Lys Ala Val Leu Ser		1295
	1300	1305
Thr Leu Val Gly Arg Tyr Gly Arg Arg Thr Lys Leu Tyr Glu Ala Ala		1310
	1315	1320
His Ser Asp Val Arg Glu Ser Leu Ala Arg Phe Ile Pro Thr Ile Gly		1325
	1330	1335
Pro Val Gln Ala Thr Thr Cys Glu Leu Tyr Glu Leu Val Glu Ala Met		1340
1345	1350	1355
Val Glu Lys Gly Gln Asp Gly Ser Ala Val Leu Glu Leu Asp Leu Cys		1360
	1365	1370
Asn Arg Asp Val Ser Arg Ile Thr Phe Phe Gln Lys Xaa Cys Asn Lys		1375
	1380	1385
Phe Thr Thr Gly Glu Thr Ile Ala His Gly Lys Val Gly Gln Gly Ile		1390
	1395	1400
Ser Ala Trp Ser Lys Thr Phe Cys Ala Leu Phe Gly Pro Trp Phe Arg		1405
	1410	1415
Ala Ile Glu Lys Glu Ile Leu Ala Leu Leu Pro Pro Asn Ile Phe Tyr		1420
1425	1430	1435
Gly Asp Ala Tyr Glu Glu Ser Val Phe Ala Ala Ala Val Ser Gly Ala		1440
	1445	1450
Gly Ser Cys Met Val Phe Glu Asn Asp Phe Ser Glu Phe Asp Ser Thr		1455
	1460	1465
		1470

Gln Asn Asn Phe Ser Leu Gly Leu Glu Cys Val Val Met Glu Glu Cys
 1475 1480 1485
 Gly Met Pro Gln Trp Leu Ile Arg Leu Tyr His Leu Val Arg Ser Ala
 1490 1495 1500
 Trp Ile Leu Gln Ala Pro Lys Glu Ser Leu Lys Gly Phe Trp Lys Lys
 1505 1510 1515 1520
 His Ser Gly Glu Pro Gly Thr Leu Leu Trp Asn Thr Val Trp Asn Met
 1525 1530 1535
 Ala Ile Ile Ala His Cys Tyr Glu Phe Arg Asp Phe Arg Val Ala Ala
 1540 1545 1550
 Phe Lys Gly Asp Asp Ser Val Val Leu Cys Ser Asp Tyr Arg Gln Ser
 1555 1560 1565
 Arg Asn Ala Ala Ala Leu Ile Ala Gly Cys Gly Leu Lys Leu Lys Val
 1570 1575 1580
 Asp Tyr Arg Pro Ile Gly Leu Tyr Ala Gly Val Val Val Ala Pro Gly
 1585 1590 1595 1600
 Leu Gly Thr Leu Pro Asp Val Val Arg Phe Ala Gly Arg Leu Ser Glu
 1605 1610 1615
 Lys Asn Trp Gly Pro Gly Pro Glu Arg Ala Glu Gln Leu Arg Leu Ala
 1620 1625 1630
 Val Cys Asp Phe Leu Arg Gly Leu Thr Asn Val Ala Gln Val Cys Val
 1635 1640 1645
 Asp Val Val Ser Arg Val Tyr Gly Val Ser Pro Gly Leu Val His Asn
 1650 1655 1660
 Leu Ile Gly Met Leu Gln Thr Ile Ala Asp Gly Lys Ala His Phe Thr
 1665 1670 1675 1680
 Glu Thr Ile Lys Pro Val Leu Asp Leu Thr Asn Ser Ile Ile Gln Arg
 1685 1690 1695
 Val Glu

<210> 92
 <211> 660
 <212> PRT
 <213> Hepatitis E Virus

<220>
 <223> Xaa = Unknown or Other at position 481
 <223> Xaa = Unknown or Other at position 542
 <223> Xaa = Unknown or Other at position 595

<400> 92
 Met Arg Pro Arg Ala Val Leu Leu Leu Phe Leu Met Phe Leu Pro Met
 1 5 10 15
 Leu Pro Ala Pro Pro Ala Gly Gln Pro Ser Gly Arg Arg Arg Gly Arg
 20 25 30
 Arg Ser Gly Gly Ala Gly Gly Gly Phe Trp Ser Asp Arg Val Asp Ser
 35 40 45
 Gln Pro Phe Ala Leu Pro Tyr Ile His Pro Thr Asn Pro Phe Ala Ala
 50 55 60
 Asp Val Val Ser Gln Pro Gly Ala Gly Thr Arg Pro Arg Gln Pro Pro
 65 70 75 80
 Arg Pro Leu Gly Ser Ala Trp Arg Asp Gln Ser Lys Arg Pro Ser Val
 85 90 95
 Ala Pro Arg Arg Arg Ser Thr Pro Ala Gly Ala Ala Pro Leu Thr Ala

			100					105				110			
Ile	Ser	Pro	Ala	Pro	Asp	Thr	Ala	Pro	Val	Pro	Asp	Val	Asp	Ser	Arg
		115					120					125			
Gly	Ala	Ile	Leu	Arg	Arg	Gln	Tyr	Asn	Leu	Ser	Thr	Ser	Pro	Leu	Thr
		130				135						140			
Ser	Ser	Val	Ala	Ser	Gly	Thr	Asn	Leu	Val	Leu	Tyr	Ala	Ala	Pro	Leu
145					150					155					160
Asn	Pro	Leu	Leu	Pro	Leu	Gln	Asp	Gly	Thr	Asn	Thr	His	Ile	Met	Ala
			165						170					175	
Thr	Glu	Ala	Ser	Asn	Tyr	Ala	Gln	Tyr	Arg	Val	Val	Arg	Ala	Thr	Ile
			180						185					190	
Arg	Tyr	Arg	Pro	Leu	Val	Pro	Asn	Ala	Val	Gly	Gly	Tyr	Ala	Ile	Ser
		195					200					205			
Ile	Ser	Phe	Trp	Pro	Gln	Thr	Thr	Thr	Thr	Pro	Thr	Ser	Val	Asp	Met
		210				215						220			
Asn	Ser	Ile	Thr	Ser	Thr	Asp	Val	Arg	Ile	Leu	Val	Gln	Pro	Gly	Ile
225					230						235				240
Ala	Ser	Glu	Leu	Val	Ile	Pro	Ser	Glu	Arg	Leu	His	Tyr	Arg	Asn	Gln
			245						250					255	
Gly	Trp	Arg	Ser	Val	Glu	Thr	Thr	Gly	Val	Ala	Glu	Glu	Glu	Ala	Thr
			260					265						270	
Ser	Gly	Leu	Val	Met	Leu	Cys	Ile	His	Gly	Ser	Pro	Val	Asn	Ser	Tyr
		275					280					285			
Thr	Asn	Thr	Pro	Tyr	Thr	Gly	Ala	Leu	Gly	Leu	Leu	Asp	Phe	Ala	Leu
		290				295						300			
Glu	Leu	Glu	Phe	Arg	Asn	Leu	Thr	Pro	Gly	Asn	Thr	Asn	Thr	Arg	Val
305					310					315					320
Ser	Arg	Tyr	Thr	Ser	Thr	Ala	Arg	His	Arg	Leu	Arg	Arg	Gly	Ala	Asp
			325						330					335	
Gly	Thr	Ala	Glu	Leu	Thr	Thr	Thr	Ala	Ala	Thr	Arg	Phe	Met	Lys	Asp
			340					345					350		
Leu	His	Phe	Thr	Gly	Thr	Asn	Gly	Val	Gly	Glu	Val	Gly	Arg	Gly	Ile
		355					360					365			
Ala	Leu	Thr	Leu	Phe	Asn	Leu	Ala	Asp	Thr	Leu	Leu	Gly	Gly	Leu	Pro
		370				375						380			
Thr	Glu	Leu	Ile	Ser	Ser	Ala	Gly	Gly	Gln	Leu	Phe	Tyr	Ser	Arg	Pro
385					390					395					400
Val	Val	Ser	Ala	Asn	Gly	Glu	Pro	Thr	Val	Lys	Leu	Tyr	Thr	Ser	Val
			405						410					415	
Glu	Asn	Ala	Gln	Gln	Asp	Lys	Gly	Ile	Thr	Ile	Pro	His	Asp	Ile	Asp
			420					425					430		
Leu	Gly	Asp	Ser	Arg	Val	Val	Ile	Gln	Asp	Tyr	Asp	Asn	Gln	His	Glu
		435						440				445			
Gln	Asp	Arg	Pro	Thr	Pro	Ser	Pro	Ala	Pro	Ser	Arg	Pro	Phe	Ser	Val
		450				455						460			
Leu	Arg	Ala	Asn	Asp	Val	Leu	Trp	Leu	Ser	Leu	Thr	Ala	Ala	Glu	Tyr
465					470					475					480
Xaa	Gln	Thr	Thr	Tyr	Gly	Ser	Ser	Thr	Asn	Pro	Met	Tyr	Val	Ser	Asp
			485						490					495	
Thr	Val	Thr	Leu	Val	Asn	Val	Ala	Thr	Gly	Ala	Gln	Ala	Val	Ala	Arg
			500					505					510		
Ser	Leu	Asp	Trp	Ser	Lys	Val	Thr	Leu	Asp	Gly	Arg	Pro	Leu	Thr	Thr
		515					520					525			
Ile	Gln	Gln	Tyr	Ser	Lys	Lys	Phe	Tyr	Val	Leu	Pro	Leu	Xaa	Gly	Lys
		530				535					540				
Leu	Ser	Phe	Trp	Glu	Ala	Gly	Thr	Thr	Lys	Ala	Gly	Tyr	Pro	Tyr	Asn
545					550					555					560

<400> 94
tggcattact actgccattg

<220>
 <223> Primer US5P45S/20

 <400> 95
 caattctgcc ttggcgaatg 20

 <210> 96
 <211> 20
 <212> DNA
 <213> Artificial Sequence

 <220>
 <223> Primer US5P296A

 <400> 96
 aggaaacacc gatgcagaac 20

 <210> 97
 <211> 20
 <212> DNA
 <213> Artificial Sequence

 <220>
 <223> Primer US5P243A/20

 <400> 97
 tccaacctcc aagcaacgac 20

 <210> 98
 <211> 199
 <212> DNA
 <213> Hepatitis E Virus

 <220>
 <223> Clone: 199con

 <400> 98
 caattctgcc ttggcgaatg ctgtggtggt tcggccggtt ctttctcgtg tgcaaaactga 60
 gattcttatt aatttgatgc aaccccgga gttggtcttc cgcctgagg tgctttggaa 120
 tcctcctatc cagcgggtta tacataatga attagagcag tactgccggg cccgggctgg 180
 tcgttgcttg gaggttga 199

 <210> 99
 <211> 25
 <212> DNA
 <213> Hepatitis E Virus

 <220>
 <223> JE orf1-s

 <400> 99
 gttctgcac ggtgttctct tagac 25

 <210> 100
 <211> 26
 <212> DNA
 <213> Hepatitis E Virus

```

<220>
<223> JE orf1-a

<400> 100
gaatcaggag atacgagggt gtgtgg                26

<210> 101
<211> 331
<212> DNA
<213> Hepatitis E Virus

<220>
<223> us2-320

<400> 101
gtttctgcatc ggtgttttct tagaccgggtc ggccgagatg ttcagcgctg gtattctgcc        60
cctacccgtg gtcctgcggc caattgccgc cgtccgcgt tgcgtgggtct cccccctgtc        120
gaccgcacct attgttttga tggattttcc cgttgtgctt ttgctgcaga gaccgggtgtg        180
gccctttact ctttgcattga cttttggcca gctgatgttg cagaggctat ggcccggcat        240
gggatgacac gcttatacgc cgcactgcac cttccccccg aggtgctgct accacccggc        300
acctaccaca caacctcgta tctctgatt c                                331

<210> 102
<211> 1186
<212> DNA
<213> Hepatitis E Virus

<220>
<223> us2-1168

<400> 102
ctcactgtgt ccatagatgt gttgggtttt gagcttacgg atatagtcca ctgccgtatg        60
gccgccccaa gccagcgaaa ggctgttctc tcaacgcttg tggggaggta cggccgtagg        120
actaaattat atgaggcggc gcattcagat gtccgtgagt ccctagcgag gtttatcccc        180
accatcgggc ctgttcgggc taccacatgt gagctgtacg agctgggtga agccatggta        240
gagaagggtc aggaaggatc tgccgtccta gagctcgacc tttgcaatcg tgacgtctcg        300
cgcatcacat ttttccaaaa ggattgcaat aagtttacaa ctggtgagac tatcgcccat        360
ggcaaggttg gccagggcac atcggcctgg agcaagacct tctgtgctct gtttggcccc        420
tggttccgcy ccattgaaaa ggaaatatgg gccctactcc cgcctaatat cttttatggc        480
gacgcctatg aggagtcagt gtttgctgcc gctgtgtccg gggcagggtc atgtatggta        540
tttgaaaatg acttctcaga gtttgacagt acccagaata atttctctct cggccttgag        600
tgtgtgggta tggaggagtg cggcatgccc caatgggtta ttaggttgta ccatctggtc        660
cggtcagocct ggattttgca ggccgcgaag gactctctta aggggttttg gaagaagcac        720
tctggtgagc ctggtaccct tctctggaac actgtctgga acatggcgat tatagcacat        780
tgctaygagt tccgtgactt tccgtgtgcc gccttcaagg gtgatgattc agtggtcctc        840
tgtagtgact accgacagrg ccgtaacgog gctgccttaa ttgcaggctg tgggctcaaa        900
ttgaagggtt attaccgccc tatcgggcta tatgctggag tgggtgggtg ccccggtttg        960
gggacactgc ccgatgtggg gcgttttgcc ggtcgggttat ctgagaagaa ttggggccct        1020
ggcccggagc gtgctgagca gctgcgtctt gctgtttgtg atttccctcg aggggtgacg        1080
aatgttgccg aggtctgtgt tgatgttgtg tcccgtgtct atggagttag ccccgggctg        1140
gtacataacc ttattggcat gctgcagacc atcgcagatg gcaagg                                1186

<210> 103
<211> 23
<212> DNA
<213> Hepatitis E Virus

```

<220>
 <223> JE hevdf2/3 s1

 <400> 103
 gttccgcttg gcgtgaccag tcc 23

 <210> 104
 <211> 23
 <212> DNA
 <213> Hepatitis E virus

 <220>
 <223> JE hevdf2/3 a1

 <400> 104
 gagtcaacat caggtacagg agc 23

 <210> 105
 <211> 130
 <212> DNA
 <213> Hepatitis E Virus

 <220>
 <223> us2-135

 <400> 105
 gttccgcttg gcgtgaccag tcccagcgcc cctccgctgc ccccgcgtc cgatctgccc 60
 cagctggggc tgcgccgctg actgccgtgt caccggctcc tgacacagct cctgtacctg 120
 atgttgactc 130

 <210> 106
 <211> 26
 <212> DNA
 <213> Hepatitis E Virus

 <220>
 <223> JE hevdf1-s1

 <400> 106
 gatgtcattt tgtgttgctg ctcacc 26

 <210> 107
 <211> 23
 <212> DNA
 <213> Hepatitis E Virus

 <220>
 <223> hev216 a1

 <400> 107
 cgtccctacag cggcatggta ctg 23

 <210> 108
 <211> 564
 <212> DNA
 <213> Hepatitis E Virus

<220>

<223> us2-563

<400> 108

tcacccatgc	cttatgttcc	ttaccctcgt	tcaacggagg	tgtatgtccg	gtctataatt	60
ggccctggcg	gctcccatc	ottgtttcca	tcagcctgct	ctactaaatc	tacctttcat	120
gctgtcccg	ttcacatctg	ggatcrgctc	atgctctttg	gtgccaccct	gracgatcag	180
gcgttctgct	gttcacggct	tatgaacttac	ctccgtggta	ttagttataa	ggtcactgtc	240
ggtgcgcttg	tcgctaataa	gggggtggaac	gcctctgagg	atgctcttac	tgcagtgtac	300
actgcggcct	atctgaccat	ctgccatcag	cgttaccttc	gcacccaggc	gatttccaag	360
ggcatgcgcc	ggttggaggt	tgagcatgct	cagaaattta	tcacaagact	ctacagctgg	420
ctatttgaga	agtcgtggcg	tgactacatc	cccggccgcc	agcttcaatt	ttatgcacaa	480
tgccgacggt	ggctttctgc	aggcttccac	ctaracccca	gggtgcttgt	ctttgatgaa	540
tcagtaccat	gccgctgtag	gacg				564

<210> 109

<211> 24

<212> DNA

<213> Hepatitis E Virus

<220>

<223> USorf2.1'

<400> 109

gtggagctag	tacaccgacc	gcag	24
------------	------------	------	----

<210> 110

<211> 678

<212> DNA

<213> Hepatitis E Virus

<220>

<223> us2-667

<400> 110

cgcttcttgg	tggtttacgc	acagaattga	tttcgtcgcc	tgggggccaa	ctgttttact	60
cccggccggt	tgtctcagcc	aatggcgagc	caacagtaaa	gttatataca	tctgttgaga	120
atgcgcagca	agacaagggc	atcaccattc	cacatgatat	agacctgggt	gactcccggt	180
tggttatcca	ggattatgat	aaccagcayg	agcaagaccg	acctactccg	tcacctgccc	240
cctctcgccc	cttctcagtt	cttcgtgcca	atgatgtttt	gtggttttcc	ctcactgccc	300
ctgagtatga	ccagactacg	tatgggtcgt	ccaccaaccc	tatgtatgtc	tctgacacag	360
ttacgcttgt	taatgtggct	actgggtgctc	aggctgttgc	ccgctccctt	gattgggtcta	420
aagttactct	ggacggccgc	ccccttacta	ccattcagca	gtattctaag	acattttatg	480
ttctcccgct	ccgcgggaag	ctgtcctttt	gggaggctgg	cacgactaag	gccgggtacc	540
cttacaatta	taatactacc	gctagtgaac	aaattttgat	tgagaatgcg	gccggccacc	600
gtgtcgctat	ttccacctat	accactagct	taggtgccgg	tcctacctcg	atctctgcgg	660
tcggtgtact	agctccac					678

<210> 111

<211> 23

<212> DNA

<213> Hepatitis E Virus

<220>

<223> hev3301s

<400> 111
 gtatgcgagc tcatccgtgg tgc 23

 <210> 112
 <211> 25
 <212> DNA
 <213> Hepatitis E Virus

 <220>
 <223> JE hev167-a1

 <400> 112
 ccaacacact atcggacaca gtgag 25

 <210> 113
 <211> 580
 <212> DNA
 <213> Hepatitis E Virus

 <220>
 <223> us2-579

 <400> 113
 gtatgcgagc tcatccgtgg tgcctacccc aaaattcaga ccacgagccg tgtgctacgg 60
 tccctgtttt ggaacgaacc ggccatcggc caaaagtggg tttttacgca ggctgctaag 120
 gctgccaacc ctggtgcgat tacggttcac gaagctcagg gtgctacttt cacggagacc 180
 acaattatag ccacggccga cgctaggggc ctcattcagt catcccgggc ccatgctata 240
 gtcgcactca cccgccatac tgagaagtgt gttatttttg atgccccggg cttgttgccg 300
 gaggtcggca tttcggatgt tattgtcaat aactttttcc ttgccgggtg agaggtcggc 360
 catcacccgc cttctgtgat acctcgcggc aatcctgatc agaacctcgg gactctacag 420
 gcctttccgc cgtcatgtca gatcagtgtc taccatcagt tggctgagga actaggtcat 480
 cgcccgccc ctgtcgcgc cgtcttgccc ccttgccctg agcttgagca gggcctgctc 540
 tatatgccac aagaactcac tgtgtccgat agtgtgttg 580

 <210> 114
 <211> 26
 <212> DNA
 <213> Hepatitis E Virus

 <220>
 <223> HEV459 s1

 <400> 114
 cagaaattta tcacaagact ctacag 26

 <210> 115
 <211> 26
 <212> DNA
 <213> Hepatitis E Virus

 <220>
 <223> HEV459 s3

 <400> 115
 ctctacagtt ggctatttga gaagtc 26

 <210> 116

```

<211> 25
<212> DNA
<213> Hepatitis E Virus

<220>
<223> JE1955a

<400> 116
ctataaagag ctgagcagaa ggcgg                25

<210> 117
<211> 734
<212> DNA
<213> Hepatitis E Virus

<220>
<223> us2-733

<400> 117
ctctacagtt ggctatttga gaagtctggc cgtgactaca tccccggccg ccagcttcaa    60
ttttatgcac aatgccgacg gtggctttct gcaggcttcc acctaraccc caggrtgctt    120
gtctttgatg aatcagtgcc atgccgttgc aggacgtttt tgaagaaggc cgcgggtaaa    180
ttctgctgtt ttatgcggtg gctggggcag gagtgtacct gcttcttggg gccagccgag    240
ggtttagttg gtgatcaagg tcatgacaac gaggcctatg aaggttctga ggtcgaccca    300
gctgagcctg cacatcttga tgtctcgggg acttatgccg tccatgggca ccagcttgag    360
gccctctata gggcacttaa tgtcccacat gatattgccg ctcgagcctc ccgactaacg    420
gctactgttg agctcgttgc tagtccggac cgottagagt gccgcactgt acttggtaat    480
aagaccttcc ggacgacggt ggttgatggc gcccatcttg aagcgaatgg ccctgaggag    540
tatgttctgt catttgacgc ctctcgccag tctatggggg ccgggtcgca cagcctcact    600
tatgagctca cccctgccgg tctgcaggta aagatttcat ctaatggtct ggattgcact    660
gccacattcc ccccyggtgg cgcccctagc gccgcgcggg gggaggtggc cgccttctgc    720
tcagctcttt atag                                734

<210> 118
<211> 22
<212> DNA
<213> Hepatitis E Virus

<220>
<223> JE 2950mex s

<400> 118
gtgtcccccg ctctggcaag tc                22

<210> 119
<211> 22
<212> DNA
<213> Hepatitis E Virus

<220>
<223> JE us2-579-a2

<400> 119
cagggttggc agccttagca gc                22

<210> 120
<211> 483

```

<212> DNA
<213> Hepatitis E Virus

<220>
<223> us2-482

<400> 120
gtgtccccgg ctctggcaag tcaaggcca tacaacaggg agatgtcgat gtggtggttg 60
tgcccacccg ggagctccgt aacagctggc gtcgocgggg ttttgccggc ttcacacctc 120
acacagcggc ccgtgttact atcggccggc gcgttgtgat tgatgaggct ccactctctc 180
cacgcacct gctgctgtta cacatgcagc gggcctcctc ggtccatctc cttggtgatc 240
caaaccagat tcctgctatt gatcttgagc atgcccggcct ggtccccgcg atccgccccg 300
agcttgccgc aacgagctgg tggcacgtta cacaccgttg cccggccgat gtgtgcgagc 360
tcatacgtgg ggccctacccc aaaattcaga ccacgagccg tgtgctacgg tccctgtttt 420
ggaacgaacc ggccatcggc caaaagttgg tttttacgca ggctgctaag gctgccaacc 480
ctg 483

<210> 121
<211> 24
<212> DNA
<213> Hepatitis E Virus

<220>
<223> JE 2600s

<400> 121
taacccaaag aggcttgagg ctgc 24

<210> 122
<211> 22
<212> DNA
<213> Hepatitis E Virus

<220>
<223> us2-482-a1

<400> 122
ccgctgtgtg aggtgtgaag gc 22

<210> 123
<211> 23
<212> DNA
<213> Hepatitis E Virus

<220>
<223> us2-482-a2

<400> 123
gacgccagct gttacggagc tcc 23

<210> 124
<211> 431
<212> DNA
<213> Hepatitis E Virus

<220>
<223> us2-430

<400> 124
 taacccaaag aggcttgagg ctgcgtaccg ggaaacttgc tcccgctcgtg gcaccgctgc 60
 ctaccgcgtt ttgggtctcg gtatatacca ggtccctggt agcctcagtt ttgatgcctg 120
 ggaacgcaat caccgccccg gcgatgagct ttacttgaca gagcccgccg cagcctgggt 180
 tgaggctaata aagccggcgc agccggcgct tactataact gaggacacgg cccgtacggc 240
 caacctggca ttagagattg atgccgccac agaggttggc cgtgcttgtg ccggctgcac 300
 catcagcccc gggattgtgc actatcagtt tacgcgcggg gtcccgggct caggcaagtc 360
 aaggtcata caacagggag atgtcgatgt ggtgggtgtg cccaccggg agctccgtaa 420
 cagctggcgt c 431

<210> 125
 <211> 22
 <212> DNA
 <213> Hepatitis E Virus

<220>
 <223> us2-orf2/3 s1

<400> 125
 cgtcgtcgat ctgccccagc tg 22

<210> 126
 <211> 25
 <212> DNA
 <213> Hepatitis E Virus

<220>
 <223> HEVConsORF2-a1

<400> 126
 cttgttcttg ytggttrtca taatc 25

<210> 127
 <211> 21
 <212> DNA
 <213> Hepatitis E Virus

<220>
 <223> us2-orf2/3 s2

<400> 127
 cgctgactgc cgtgtcaccg g 21

<210> 128
 <211> 25
 <212> DNA
 <213> Hepatitis E Virus

<220>
 <223> HEVConsORF2-a2

<400> 128
 gttcrtgytg gtttrcataa tcttg 25

<210> 129
 <211> 1020

<212> DNA

<213> Hepatitis E Virus

<220>

<223> us2-1019

<400> 129

cgctgactgc	cgtgtcaccg	gctcctgaca	cagccccctgt	acctgatgtt	gactcacgtg	60
gtgctattct	gcgccggcag	tacaatttgt	ccacgtcccc	gctcacgtca	tctgtcgott	120
cgggtactaa	tttggtcctc	tatgctgccc	cgtgaatcc	cctcttgctt	ctccaggatg	180
gtaccaaac	tcatattatg	gctactgagg	catccaatta	tgcccagtat	cgggttggtc	240
gagctacaat	ccgttatcgc	ccgctggtgc	cgaatgccgt	tgggtggctat	gccatttcca	300
tttctttctg	gccccaaact	acaactaccc	ctacttctgt	cgatatgaat	tctattactt	360
ccacygatgt	taggatTTTT	gttcagcccc	gtattgcctc	cgagctagtc	atccccagtg	420
agcgcttca	ttaccgtaat	caaggctggc	gctctgttga	gaccacgggt	gtggctgagg	480
aggaggctac	ttccggtctg	gtaatgcttt	gcattcatgg	ctctcctggt	aattcctaca	540
ctaatacacc	ttacactggt	gcgctggggc	ttcttgattt	tgcactagag	cttgaattta	600
ggaatttgac	acccgggaac	accaaccccc	gtgtttcccc	gtataccagc	acagcccgcc	660
accggtgcg	cogtggtgct	gatgggactg	ctgagcttac	taccacagca	gccacacgtt	720
tcatgaagga	cctgcacttc	gctggcacga	atggcggttg	tgagggtggg	cgtggtatcg	780
ccctgacact	gttcaatctc	gctgatacgc	ttctcgccgg	tttaccgaca	gaattgattt	840
cgctcggtgg	gggccaaactg	ttttactccc	gcccgggttg	ctcagccaat	ggcgagccaa	900
cagtaaagtt	atatacatct	gttgagaatg	cgcagcaaga	caagggcata	accattccac	960
atgatataga	cctgggtgac	tcccgtgtgg	ttatccagga	ttatgataac	cagcaygaac	1020

<210> 130

<211> 24

<212> DNA

<213> Hepatitis E Virus

<220>

<223> us2 330s1

<400> 130

cagctgatgt	tgcagaggct	atgg	24
------------	------------	------	----

<210> 131

<211> 24

<212> DNA

<213> Hepatitis E Virus

<220>

<223> us2 563a1

<400> 131

gcaggctgat	ggaaacaagg	atgg	24
------------	------------	------	----

<210> 132

<211> 407

<212> DNA

<213> Hepatitis E Virus

<220>

<223> us2-406

<400> 132

cagctgatgt	tgcagaggct	atggccccgc	atgggatgac	acgcttatac	gccgcactgc	60
------------	------------	------------	------------	------------	------------	----

```

accttccccc cgaggtgctg ctaccacccg gcacctacca cacaacctcg tacctcttga 120
ttcacgatgg caaccgcgct gttgtaactt acgagggcga tactagtgcg ggctataatc 180
atgatgtctc catacttcgt gcatggatcc gtactactaa aatagttggg gaccatccat 240
tggtcataga gcgagtgcgg gccattgggt gtcattttgt gctgctgctc accgcagccc 300
ctgaaccgtc acctatgcct tatgttcctt accctcgttc aacggagggt tatgtccggt 360
ctatatttgg ccctggcggc tcccacatct tgtttccatc agcctgc 407

```

```

<210> 133
<211> 22
<212> DNA
<213> Hepatitis E Virus

```

```

<220>
<223> us2-579 s1

```

```

<400> 133
cagaccacga gccgtgtgct ac 22

```

```

<210> 134
<211> 23
<212> DNA
<213> Hepatitis E Virus

```

```

<220>
<223> us2-1168 a1

```

```

<400> 134
ccacaagcgt tgagagaaca gcc 23

```

```

<210> 135
<211> 22
<212> DNA
<213> Hepatitis E Virus

```

```

<220>
<223> us2-579 s2

```

```

<400> 135
gctgctaagg ctgccaaacc tg 22

```

```

<210> 136
<211> 547
<212> DNA
<213> Hepatitis E Virus

```

```

<220>
<223> us2-579wb

```

```

<400> 136
gctgctaagg ctgccaaacc tgggtgcgatt acggttcacg aagctcaggg tgctactttc 60
acggagacca caattatagc cacggccgac gctaggggac tcattcagtc atccccggcc 120
catgctatag tcgcactcac ccgccatact gagaagtgtg ttatttttga tgccccggcc 180
ttgttgccgg aggtcggcat ttccgatgtt attgtcaata actttttcct tgccggtgga 240
gaggtcggcc atcaccgccc ttctgtgata cctcgcggca atcctgatca gaacctcggg 300
actctacagg cctttccgcc gtcatgtcag atcagtgcct accatcagtt ggctgaggaa 360
ctaggtcatc gcccgcccc tgtcgcccc gtcttgcccc cttgccctga gcttgagcag 420
ggcctgctct atatgccaca agaacttact gtgtccgata gcgtgctggt ttttgagctt 480

```

acggatatag tccactgccg tatggccgcc ccaagccagc gaaaggtgt tctctcaacg 540
cttgtgg 547

<210> 137
<211> 24
<212> DNA
<213> Hepatitis E Virus

<220>
<223> us2-733s1

<400> 137
cacagcctca cttatgagct cacc 24

<210> 138
<211> 23
<212> DNA
<213> Hepatitis E Virus

<220>
<223> us2-430a1

<400> 138
cggatgattgc gttcccaggc atc 23

<210> 139
<211> 26
<212> DNA
<213> Hepatitis E Virus

<220>
<223> us2-733s2

<400> 139
ctgcaggtaa agatttcac tc taatgg 26

<210> 140
<211> 24
<212> DNA
<213> Hepatitis E Virus

<220>
<223> us2-430a2

<400> 140
ccaggcatca aaactgaggc taac 24

<210> 141
<211> 903
<212> DNA
<213> Hepatitis E Virus

<220>
<223> us2-851

<400> 141
ctgcaggtaa agatttcac tc taatggtctg gattgcactg ccacattccc cccygggtggc 60

```

gcccctagcg ccgcgcgggg ggaggtggcs gccttctgca gtgctcttta tagatacaat 120
aggttcaccc agcggcattc gctgacaggc ggactatggc tacatcctga ggggctgctg 180
ggatatcttc cccattcttc ccctgggcat atttgggagt ctgctaacc cttttgcggg 240
gaggggactt tgtatacccg aacctgggtc acctctgggt tttctagtga tttctcccc 300
cctgaggcgg ccgctectgc ttcggctgcc gcccgggggt tgccctacc tactccacct 360
gttagtgata tctgggtggt accaccgccc tcagaggaat ctcatgttga tgcggcatct 420
gtaccctctg ttcttgagcc tgctggattg accagcccta ttgtgcttac ccccccccc 480
ccccctctc ccgtgcgtaa gccggcaaca tccccgcctc ccgcactcg ccgtctcctt 540
tacacctacc ccgacggcgc caaggtgtat gcgggggtcat tgtktagagc agactgtgat 600
tggttagtca atgcctcaaa ccctggccat cgcggcggtg gtggcctctg ccatgctttt 660
tatcaacgtt tcccagaagc gttctactcg actgaattca tcatgcgcga gggccttgca 720
gcatacactt taaccccgcg ccctattatc catgcagtgg ctcccgacta taggggtgag 780
caaaacccga agaggcttga ggcagcgtac cgggaaactt gctcccgtcg tggcaccgct 840
gcctaccgcg ttttgggctc gggatatatac caggtccctg ttagcctcag ttttgatgcc 900
tgg 903

```

```

<210> 142
<211> 24
<212> DNA
<213> Hepatitis E Virus

```

```

<220>
<223> us2-1168s1

```

```

<400> 142
gcaggtctgt gttgatgttg tgtc 24

```

```

<210> 143
<211> 21
<212> DNA
<213> Hepatitis E Virus

```

```

<220>
<223> us2-dforf2/3 a2

```

```

<400> 143
ccggtgacac ggcagtcagc g 21

```

```

<210> 144
<211> 25
<212> DNA
<213> Hepatitis E Virus

```

```

<220>
<223> us2-1168s2

```

```

<400> 144
gatgttgtgt ccggtgtcta tggag 25

```

```

<210> 145
<211> 22
<212> DNA
<213> Hepatitis E Virus

```

```

<220>
<223> us2 dforf2/3 a3

```

<400> 145
 cagctggggc agatcgacga cg 22

<210> 146
 <211> 503
 <212> DNA
 <213> Hepatitis E Virus

<220>
 <223> us2-502

<400> 146
 gatgttggtg cccgtgtcta tggagtttag cccgggctgg tacataacct tattggcatg 60
 ctgcagacca ttgctgatgg caaggcccac tttacagara atattaaacc tgtgcttgac 120
 cttacaaatt ccatcataca acgggtggaa tgaataacat gtcttttgca tcgcccattg 180
 gatcaccatg cggcctaggg ctgttctgtt gttgctcttc gtgcttttgc ctatgctgcc 240
 cgcgccaccg gccggccagc cgtctggccg ccgtcgtggg cggcgcagcg gcggtgccgg 300
 cggtggtttc tggggtgaca gggttgattc tcagcccttc gccctcccct atattcatcc 360
 aaccaacccc ttccgcccg atgtcgtttc acaaccggg gctggaaact gccctcgaca 420
 gccgccccgc ccccttgggt ccgcttgggc tgaccagtc cagcggccct ccgctgcccc 480
 ccgtcgtcga tctgcccag ctg 503

<210> 147
 <211> 24
 <212> DNA
 <213> Hepatitis E Virus

<220>
 <223> HEVConsORF1-s1

<400> 147
 ctggcatyac tactgcyatt gagc 24

<210> 148
 <211> 23
 <212> DNA
 <213> Hepatitis E Virus

<220>
 <223> HEVConsORF1-a1

<400> 148
 ccatcrarrc agtaagtgcg gtc 23

<210> 149
 <211> 418
 <212> DNA
 <213> Hepatitis E Virus

<220>
 <223> us2-orf1

<400> 149
 ctggcattac tactgtatt gagcaggtg ctctggctgc ggctaattcc gccttggcga 60
 atgtgtggt ggttcggccg tttctttctc gtgtgcaaac tgagattctt attaatgtga 120
 tgcaaccccg gcagttggc ttccgccctg aggtgctttg gaatcactct atccagcggg 180
 ttatacataa tgaattagag cagtactgcc gggcccgggc tggctggtgt ttggaggttg 240

gagccacccc	gaggtccatt	aatgacaacc	ctaattgtctt	gcataggtgt	tttcttagac	300
cggtcggcgc	agatgttcag	cgctgggtatt	ctgcccctac	ccgtgggtcct	gcggccaatt	360
gccgcgcgtc	cgcggttgct	gggtctcccc	ctgtcgaccg	cacttactgt	tttgatgg	418

<210> 150
 <211> 24
 <212> DNA
 <213> Hepatitis E Virus

<220>
 <223> HEVConsORF2-s1

<400> 150	
gacagaattt	atttcgtcgg ctgg 24

<210> 151
 <211> 197
 <212> DNA
 <213> Hepatitis E Virus

<220>
 <223> us2-orf2

<400> 151	
gacagaattg	atttcgtcgg ctgggggcca actgttttac tcccgcccg ttgtctcagc 60
caatggcgag	ccaacagtaa agttatatac atctgttgag aatgcgcagc aagacaaggg 120
catcaccatt	ccacatgata tagacctggg tgactcccg gtggttatcc aggattatga 180
taaccagcay	gagcaag 197

<210> 152
 <211> 22
 <212> DNA
 <213> Hepatitis E Virus

<220>
 <223> HEVConsORF2-s2

<400> 152	
gtygtctcrg	ccaatggcga gc 22

<210> 153
 <211> 901
 <212> DNA
 <213> Hepatitis E Virus

<220>
 <223> us2-3p

<400> 153	
gttgtctcag	ccaatggcga gccaacagta aagttatata catctgttga gaatgcgcag 60
caagacaagg	gcatcaccat tccacatgat atagacctgg gtgactcccg tgtgggtatc 120
caggattatg	ataaccagca ygagcaagac cgacctactc cgtcacctgc cccctctcgc 180
cccttctcag	ttcttcgtgc caatgatgtt ttgtggcttt cctcactgc cgtcgagtat 240
gaccagacta	cgtatgggtc gtccaccaac cctatgtatg tctctgacac agttacgctt 300
gttaattgtg	ctactgggtc tcaggtgtt gcccgctccc ttgattggtc taaagttact 360
ctggacggcc	gcccccttac taccattcag cagtattcta agacatttta tgttctcccg 420
ctccgcggga	agctgtcctt ttgggaggct ggcacgacta aggcgggcta cccttacaat 480

tataatacta	ccgctagtga	ccaaattttg	attgagaatg	cggccggcca	ccgtgtcgct	540
atttccacct	ataccactag	cttaggtgcc	ggtcctacct	cgatctctgc	ggtcggcgta	600
ctggctccac	actctgccct	tgccgttctt	gaggatacta	ttgattaccc	cgcccggtgcc	660
catacttttg	atgatttttg	cccggagtgc	cgtaccctag	gtttgcaggg	ttgtgcattc	720
cagtctacta	ttgctgagct	ccagcgttta	aaaatgaagg	taggtaaaac	ccgggagtct	780
taattaattc	cttctgtgcc	cccttcgtag	tttctttcgc	ttttatttct	tattttctgt	840
ttcgcgctc	cctggaaaaa	aaaaaaaaaa	aaaaaaaaaa	agtactagtc	gacgcgtggc	900
c						901

<210> 154

<211> 27

<212> DNA

<213> Hepatitis E Virus

<220>

<223> us2-gap s1

<400> 154

tatagataac aatagggttca cccagcgc

27

<210> 155

<211> 25

<212> DNA

<213> Hepatitis E Virus

<220>

<223> us2-gap a1

<400> 155

attcagtcga gtagaacgct totgg

25

<210> 156

<211> 23

<212> DNA

<213> Hepatitis E Virus

<220>

<223> us2-gap s2

<400> 156

cggactatgg ctacatcctg agg

23

<210> 157

<211> 26

<212> DNA

<213> Hepatitis E Virus

<220>

<223> us2-gap a2

<400> 157

ttgactaacc aatcacagtc tgactc

26

<210> 158

<211> 462

<212> DNA

<213> Artificial Sequence

<220>

<223> 13906-gap

<400> 158

cggactatgg	ctacatcctg	aggggctgct	gggtatcttc	ccccattct	ccctgggca	60
tatttgggag	tctgctaacc	ccttttgcg	tgaggggact	ttgtatacc	gaacctggc	120
aacctctggt	ttttctagt	atttctccc	ccctgaggcg	gctgctcctg	cttcggctgc	180
cgccccgggg	ttgcctacc	ctactccacc	tgtagtgat	atctgggtgt	taccaccgc	240
ctcagaggaa	tctcatgtt	atgcggcatc	tgtacctct	gttctgagc	ctgctggatt	300
gaccagccct	attgtgctta	ccccccccc	ccccctcct	ccgtgcgta	agcggcaac	360
atccccgct	ccccgactc	gcgtctcct	ttacacctac	ccgacggcg	ccaagggtga	420
tcgggggtca	ttgtttgagt	cagactgtga	ttggttagtc	aa		462

<210> 159

<211> 21

<212> DNA

<213> Hepatitis E Virus

<220>

<223> us-575a

<400> 159

gccgggtggt	agcagcacct	c	21
------------	------------	---	----

<210> 160

<211> 24

<212> DNA

<213> Hepatitis E Virus

<220>

<223> us-426s

<400> 160

cgttgtgctt	ttgctgcaga	gacc	24
------------	------------	------	----

<210> 161

<211> 22

<212> DNA

<213> Hepatitis E Virus

<220>

<223> us-84a

<400> 161

gaaacggccg	aaccaccaca	gc	22
------------	------------	----	----

<210> 162

<211> 24

<212> DNA

<213> Hepatitis E Virus

<220>

<223> us-484s

<400> 162

cagctgatgt	tcagaggct	atgg	24
------------	-----------	------	----

<210> 163
 <211> 22
 <212> DNA
 <213> Hepatitis E Virus

<220>
 <223> us-78a

<400> 163
 gccgaaccac cacagcattc gc

22

<210> 164
 <211> 7277
 <212> DNA
 <213> Hepatitis E Virus

<220>
 <223> us2fu11

<400> 164
 tcgacagggg gcagaccacg tatgtggtcg atgccatgga ggcccatcag ttcattaagg 60
 ctccctggcat tactactgct attgagcagg ctgctctggc tgcggctaata tccgccttgg 120
 cgaatgctgt ggtggttcgg ccgtttcttt ctgctgtgca aactgagatt cttattaatt 180
 tgatgcaacc ccggcagttg gtcttcggcc ctgaggtgct ttggaatcat cctatccagg 240
 ggggtataca taatgaatta gagcagtact gccgggcccg ggctggtcgt tgtttggagg 300
 ttggagccca ccgaggtcc attaatgaca accctaattgt cttgcatagg tgttttctta 360
 gaccggtcgg ccgagatggt cagcgtggtt attctgcccc taccogtggg cctgcggcca 420
 attgcccgg ctcccggttg cgtggtctcc ccctgtcga ccgcacctat tgttttgatg 480
 gattttcccg ttgtgctttt gctgcagaga ccggtgtggc cctttactct ttgcatgacc 540
 tttggccagc tgatgttgca gaggtatgg ccgcctatgg gatgacaagc ttatacggcg 600
 cactgcacct tccccccgag gtgctgtctac caccggcac ctaccacaca acctcgtacc 660
 tcttgattca cgatggcaac cgcgctgttg taacttacga gggcgatact agtgcgggct 720
 ataactcatga tgtctccata cttcgtgcat ggatccgtac tactaaaata gttggtgacc 780
 atccattggt catagagcga gtgcggggcca ttgggtgtca ttttgtgctg ctgctcaccg 840
 cagccctga accgtcacct atgccttatg ttccctaccc tcgttcaacg gaggtgtatg 900
 tccggtctat atttgccctt ggcggctccc catccttgtt tccatcagcc tgcctacta 960
 aatctacctt tcatgtgtc ccggttcaca tctgggacxr gctcatgctc tttggtgcca 1020
 cccgracga tcaggcgttc tgctgttcac ggcttatgac ttacctccgt ggtattagtt 1080
 ataaggtcac tgcgggtgcy cttgtcgcta atgaggggtg gaacgcctct gaggatgctc 1140
 ttactgcagt gatcactgcy gctatctga ccatctgcca tcagcggtac cttcgcaccc 1200
 aggcgatttc caaggcagtg cgcgggttgg aggttgagca tgcctagaaa tttatcacia 1260
 gactctacag ctggctattt gagaagtctg gccgtgacta catccccggc cggcagcttc 1320
 aatttttatgc acaatgccga cgggtgcttt ctgcaggctt ccacctarac ccaggrtgc 1380
 ttgtctttga tgaatcagtg ccatgccgtt gcaggacgtt tttgaagaag gtcgcgggta 1440
 aattctgctg ttttatgcgg tggctggggc aggagtgtac ctgcttcttg gagccagccg 1500
 agggtttagt tgggtgatcaa ggtcatgaca acgaggccta tgaaggttct gaggtcgacc 1560
 cagctgagcc tgcacatctt gatgtctcgg ggacttatgc cgtccatggg caccagcttg 1620
 aggcctctta tagggcactt aatgtccac atgatattgc cgtccgagcc tcccactaa 1680
 cggctactgt tgagctcgtt gctagtccgg accgcttaga gtgccgact gtacttggtta 1740
 ataagacctt ccggacgacg gtgggtgatg gcgccatct tgaagcgaat ggccctgagg 1800
 agtatgttct gtcatttgac gctctcggc agtctatggg ggccgggtcg cacagcctca 1860
 cttatgagct caccctgccc ggtctgcagg taaagatttc atctaattgt ctggattgca 1920
 ctgccacatt ccccccgggt ggccgcccta gcgccgcgc gggggagggt gcsgccttct 1980
 gcagtgtctt ttatagatac aataggttca ccagcggca ttccgtgaca ggccgactat 2040
 ggctacatcc tgaggggctg ctgggtatct tccccatt ctccctggg catatttggg 2100
 agtctgctaa ccccttttgc ggtgagggga ctttgtatac ccgaacctgg tcaacctctg 2160

gtttttctag	tgattttctcc	ccccctgagg	cgcccgctcc	tgcttcgget	gccgccccgg	2220
ggttgcacct	ccctactcca	cctgttagtg	atatctgggt	gttaccaccg	ccctcagagg	2280
aatctcatgt	tgatgcggca	tctgtaccct	ctgttcctga	gctgctgga	ttgaccagcc	2340
ctattgtgct	tacccccccc	ccccccctc	ctcccgctcg	taagccggca	acatccccgc	2400
ctccccgcac	tgcgcgtctc	ctttacacct	accccgacgg	cgccaagggt	tatgcggggg	2460
cattgtktga	gtcagactgt	gatttggttag	tcaatgcctc	aaaccctggc	catcgccccg	2520
ggggtggcct	ctgccatgct	ttttatcaac	gtttcccaga	agcgttctac	tcgactgaat	2580
tcatcatgcg	cgaggggcctt	gcagcataca	ctttaacccc	gcgccttatt	atccatgcag	2640
tggtctccga	ctatagggtt	gagcaaaacc	cgaagaggct	tgaggcagcg	taccgggaaa	2700
cttgcctccg	togtggcacc	gctgcctacc	cgcttttggg	ctcgggtata	taccaggtcc	2760
ctgttagcct	cagttttgat	gcttgggaac	gcaatcaccc	ccccggcgat	gagctttact	2820
tgacagagcc	cgccgcagcc	tggttttgagg	ctaataagcc	ggcgagcccg	gcgcttacta	2880
taactgagga	cacgggccgt	acggccaacc	tggcattaga	gattgatgcc	gccacagagg	2940
ttggcgcgtg	ttgtgccggc	tgcaccatca	gccccgggat	tgtgcactat	cagtttaccg	3000
ccgggggtccc	gggctcaggg	aagtcaagg	ccatacaaca	gggagatgtc	gatgtgggtg	3060
ttgtgcccac	ccgggagctc	cgtaacagct	ggcgtcgccg	gggttttgcg	gccttcacac	3120
ctcacacagc	ggcccgctgt	actatcgcc	gcgcgttgt	gattgatgag	gctccatctc	3180
tcccaccgca	cctgctgctg	ttacacatgc	agcgggcctc	ctcggctcat	ctccttgggtg	3240
atccaaacca	gattcctgct	attgattttg	agcatgccgg	cctggtcccc	gcgatccgcc	3300
ccgagcttgc	gccaacgagc	tggtggcacc	ttacacaccg	ttgcccgcc	gatgtgtgcg	3360
agctcatacg	tggggcctac	cccaaaattc	agaccacgag	ccgtgtgcta	cggtccctgt	3420
tttggaacga	acccggccatc	ggccaaaagt	tggtttttac	gcaggctgct	aaggctgcca	3480
acccgtgtgc	gattacgggt	cacgaagctc	agggtgtctac	tttcacggag	accacaatta	3540
tagccacggc	cgacgcctagg	ggcctcattc	agtcaccccg	ggcccatgct	atagtcgcac	3600
tcaccgcgca	tactgagaag	tgtgttattt	tggatgcccc	cggcttgttg	cgcgaggctg	3660
gcattttcgg	tgttattgtc	aataactttt	tccttgccgg	tggagagggt	ggccatcacc	3720
gcccctctgt	gatacctcgc	ggcaatcctg	atcagaacct	cgggactcta	caggcccttc	3780
cgccgtcatg	tcagatcagt	gcttaccatc	agttggctga	ggaactaggt	catcgccccg	3840
ccccctgtgc	cgccgtcttg	cccccttgcc	ctgagcttga	gcagggcctg	ctctatatgc	3900
cacaagaact	tactgtgtcc	gatagcgtgc	tggtttttga	gcttacggat	atagtccact	3960
gccttatggc	cgccccaagc	cagcgaaaag	ctgttctctc	aacgcttgtg	gggaggtacg	4020
gcccgtaggc	taaattatat	gaggcggcgc	attcagatgt	ccgtgagtc	ctagcgagggt	4080
ttatccccac	catcgggcct	gttcgggcta	ccacatgtga	gctgtacgag	ctggttgaag	4140
ccatggtaga	gaagggtcag	gacggatctg	ccgtccctaga	gctcgacctt	tgcaatcgtg	4200
acgtctcgcg	catcacattt	ttccaaaagg	attgcaataa	gtttacaact	ggtgagacta	4260
tcgccccatg	caagggttgg	cagggcatac	cggccgtggg	caagaccttc	tgtgctctgt	4320
ttggcccggt	gttcgcgcgc	attgaaaagg	aaataattgg	cctactcccc	cctaatactc	4380
tttatggcga	cgccatagag	gagtcagtgt	ttgctgccc	tgtgtccggg	gcagggtcat	4440
gtatggtatt	tgaaaatgac	ttctcagagt	ttgacagtac	ccagaataat	ttctctctcg	4500
gccttgagtg	tgtggttatg	gaggagtgcg	gcatgcccc	atggttaatt	aggttgtacc	4560
atctggtccg	gtcagcctgg	attttgcagg	cgccgaagg	gtctcttaag	gggttttgg	4620
agaagcactc	tggtgagcct	ggtacccttc	tctggaacac	tgtctggaac	atggcgatta	4680
tagcacattg	ctaygagttc	cgtgactttc	gtgttgccgc	cttcaagggt	gatgattcag	4740
tggtcctctg	tagtgactac	cgacagrgcc	gtaacgcggc	tgcccttaatt	gcaggctgtg	4800
ggctcaaaatt	gaaggttgat	taccgcccta	tcgggctata	tgctggagtg	gtggtggccc	4860
ccggtttggg	gacactgccc	gatgtgggtg	gttttgccgg	tcgggttatct	gagaagaatt	4920
ggggccctgg	cccggagcgt	gctgagcagc	tgctgtctgc	tgtttgtgat	ttccttcgag	4980
ggttgacgaa	tgttgccgag	gtctgtgttg	atgttgtgtc	ccgtgtctat	ggagttagcc	5040
ccgggctggg	acataacctt	attggcatgc	tgagaccat	tgctgatggc	aaggccact	5100
ttacagaaaa	tattaaacct	gtgcttgacc	ttacaaaatc	catcatacaa	cgggtgggaat	5160
gaataacatg	tcttttgcac	cgcccatggg	atcaccatgc	gccctagggc	tgttctgttg	5220
ttgctcttcg	tgccttttgc	tatgctgccc	gcgccaccgg	ccggccagcc	gtctggccgc	5280
cgtcgtgggc	ggcgagcggg	cgggtgcggc	ggtggtttct	ggggtgacag	ggttgattct	5340
cagcccttcg	ccctccccta	tattcatcca	accaaccctc	tcgcccgcga	tgtcgtttca	5400
caaccggggg	ctggaactcg	ccctcgacag	ccgcccggcc	cccttgggtc	cgcttggcgt	5460
gaccagtcct	agcgccctc	cgctgcccc	cgctcgtgat	ctgcccagc	tggggtcgcg	5520
ccgctgactg	ccgtgtcacc	ggctcctgac	acagccctg	tacctgatgt	tgactcacgt	5580

```

ggtgctattc tgcgcogggca gtacaatttg tccacgtccc cgtcacgctc atctgtcgct 5640
tcgggtacta atttggtoct ctatgctgcc ccgctgaatc cctcttgcc tctccaggat 5700
ggtaccaaca ctcattatgt ggctactgag gcatccaatt atgcccagta tcgggttggt 5760
cgagctacaa tccgttatcg cccgctgggtg ccgaatgccg ttggtggcta tgccatttcc 5820
atctctttct ggccccaacac tacaactacc cctacttctg tcgatatgaa ttctattact 5880
tcacacygatg ttaggatttt gggttcagccc ggtattgccc ccgagctagt catccccagt 5940
gagcgccctc attaccgtaa tcaaggctgg cgtctgtgtg agaccacggg tgtggctgag 6000
gaggaggcta cttccggctc ggtaatgctt tgcattcatg gctctcctgt taattcctac 6060
actaatacac cttacactgg tgcgctgggg cttcttgatt ttgcactaga gcttgaattt 6120
aggaatttga caccogggaa caccaacacc cgtgtttccc ggtataccag cacagcccg 6180
cacoggtgc gccgtggtgc tgatgggact gctgagctta ctaccacagc agccacacgt 6240
ttcatgaagg acctgcactt cgtggtgacg aatggcggtg gtgagggtgg tcgtggtatc 6300
gcccgtgacac tgttcaatct cgtgatacgc cttctcggtg gtttaccgac agaattgatt 6360
tcgtcggtcg ggggccaact gttttactcc cggccggttg tctcagccaa tggcgagcca 6420
acagtaaagt tatatacatc tgttgagaat gcgcagcaag acaaggcgat caccattcca 6480
catgatatag acctgggtga ctcccggtg gttatccagg attatgataa ccagcaygag 6540
caagaccgac ctactccgct acctgcccc tctcgccct tctcagttct tcgtgccaat 6600
gatgttttgt ggctttccct cactgcccgt gactatgacc agactacgta tgggtcgctc 6660
accaacccta tgtatgtctc tgacacagtt acgcttggtt atgtggctac tgggtgctcag 6720
gctgttgccc gctcccttga ttgggtctaaa gttactctgg acggccgccc cttactacc 6780
attcagcagt attctaagac attttatgtt ctcccgctcc gcgggaagct gtccttttgg 6840
gaggtgggca cgactaaggc cggctaccct tacaattata atactaccgc tagtgaccaa 6900
atgttgattg agaatgcggc cggccaccgt gtcgctattt ccacotatac cactagctta 6960
ggtgcccgtc ctacctcgat ctctgcggtc ggcgactggt ctcacacactc tgcccttgcc 7020
gttcttgagg atactattga ttaccccgcc cgtgcccata cttttgatga tttttgccc 7080
gagtgccgta ccctaggttt gcaggggtgt gcattccagt ctactattgc tgagctccag 7140
cgttataaaa tgaaggtagg taaaaccggg gactottaat taattccttc tgtgccccct 7200
tcgtagtttc tttcgctttt atttcttatt tctgctttcc gcgctccctg gaaaaaaaaa 7260
aaaaaaaaaa aaaaaaaa 7277

```

<210> 165

<211> 7277

<212> DNA

<213> Hepatitis E Virus

<220>

<223> us2fu11

<221> CDS

<222> (36) ... (5159)

<223> orf1

<221> CDS

<222> (5197) ... (7176)

<223> orf2

<223> orf3 at positions 5159-5527

<223> Xaa = Unknown or Other at position 322

<223> Xaa = Unknown or Other at position 331

<223> Xaa = Unknown or Other at position 445

<223> Xaa = Unknown or Other at position 448

<223> Xaa = Unknown or Other at position 634	
<223> Xaa = Unknown or Other at position 646	
<223> Xaa = Unknown or Other at position 811	
<223> Xaa = Unknown or Other at position 1553	
<223> Xaa = Unknown or Other at position 1578	
<223> Xaa = Unknown or Other at position 1691	
<223> Xaa = Unknown or Other at position 1792	
<223> Xaa = Unknown or Other at position 1938	
<223> Xaa = Unknown or Other at position 2155	
<400> 165	
tcgacagggg gcagaccacg tatgtggtcg atgcc atg gag gcc cat cag ttc	53
Met Glu Ala His Gln Phe	
1 5	
att aag gct cct ggc att act act gct att gag cag gct gct ctg gct	101
Ile Lys Ala Pro Gly Ile Thr Thr Ala Ile Glu Gln Ala Ala Leu Ala	
10 15 20	
gcg gct aat tcc gcc ttg gcg aat gct gtg gtg gtt cgg ccg ttt ctt	149
Ala Ala Asn Ser Ala Leu Ala Asn Ala Val Val Val Arg Pro Phe Leu	
25 30 35	
tct cgt gtg caa act gag att ctt att aat ttg atg caa ccc cgg cag	197
Ser Arg Val Gln Thr Glu Ile Leu Ile Asn Leu Met Gln Pro Arg Gln	
40 45 50	
ttg gtc ttc cgc cct gag gtg ctt tgg aat cat cct atc cag cgg gtt	245
Leu Val Phe Arg Pro Glu Val Leu Trp Asn His Pro Ile Gln Arg Val	
55 60 65 70	
ata cat aat gaa tta gag cag tac tgc cgg gcc cgg gct ggt cgt tgt	293
Ile His Asn Glu Leu Glu Gln Tyr Cys Arg Ala Arg Ala Gly Arg Cys	
75 80 85	
ttg gag gtt gga gcc cac ccg agg tcc att aat gac aac cct aat gtc	341
Leu Glu Val Gly Ala His Pro Arg Ser Ile Asn Asp Asn Pro Asn Val	
90 95 100	
ttg cat agg tgt ttt ott aga ccg gtc ggc cga gat gtt cag cgc tgg	389
Leu His Arg Cys Phe Leu Arg Pro Val Gly Arg Asp Val Gln Arg Trp	
105 110 115	
tat tct gcc cct acc cgt ggt cct gcg gcc aat tgc cgc cgc tcc gcg	437
Tyr Ser Ala Pro Thr Arg Gly Pro Ala Ala Asn Cys Arg Arg Ser Ala	
120 125 130	
ttg cgt ggt ctc ccc cct gtc gac cgc acc tat tgt ttt gat gga ttt	485

Leu	Arg	Gly	Leu	Pro	Pro	Val	Asp	Arg	Thr	Tyr	Cys	Phe	Asp	Gly	Phe	
135					140					145					150	
tcc	cgt	tgt	gct	ttt	gct	gca	gag	acc	ggg	gtg	gcc	ctt	tac	tct	ttg	533
Ser	Arg	Cys	Ala	Phe	Ala	Ala	Glu	Thr	Gly	Val	Ala	Leu	Tyr	Ser	Leu	
			155						160					165		
cat	gac	ctt	tgg	cca	gct	gat	gtt	gca	gag	gct	atg	gcc	cgc	cat	ggg	581
His	Asp	Leu	Trp	Pro	Ala	Asp	Val	Ala	Glu	Ala	Met	Ala	Arg	His	Gly	
			170					175					180			
atg	aca	cgc	tta	tac	gcc	gca	ctg	cac	ctt	ccc	ccc	gag	gtg	ctg	cta	629
Met	Thr	Arg	Leu	Tyr	Ala	Ala	Leu	His	Leu	Pro	Pro	Glu	Val	Leu	Leu	
		185					190					195				
cca	ccc	ggc	acc	tac	cac	aca	acc	tcg	tac	ctc	ttg	att	cac	gat	ggc	677
Pro	Pro	Gly	Thr	Tyr	His	Thr	Thr	Ser	Tyr	Leu	Leu	Ile	His	Asp	Gly	
	200					205					210					
aac	cgc	gct	gtt	gta	act	tac	gag	ggc	gat	act	agt	gcg	ggc	tat	aat	725
Asn	Arg	Ala	Val	Val	Thr	Tyr	Glu	Gly	Asp	Thr	Ser	Ala	Gly	Tyr	Asn	
215					220					225					230	
cat	gat	gtc	tcc	ata	ctt	cgt	gca	tgg	atc	cgt	act	act	aaa	ata	gtt	773
His	Asp	Val	Ser	Ile	Leu	Arg	Ala	Trp	Ile	Arg	Thr	Thr	Lys	Ile	Val	
			235					240						245		
ggg	gac	cat	cca	ttg	gtc	ata	gag	cga	gtg	cgg	gcc	att	ggg	tgt	cat	821
Gly	Asp	His	Pro	Leu	Val	Ile	Glu	Arg	Val	Arg	Ala	Ile	Gly	Cys	His	
			250					255					260			
ttt	gtg	ctg	ctg	ctc	acc	gca	gcc	cct	gaa	ccg	tca	cct	atg	cct	tat	869
Phe	Val	Leu	Leu	Leu	Thr	Ala	Ala	Pro	Glu	Pro	Ser	Pro	Met	Pro	Tyr	
		265					270					275				
gtt	ccc	tac	cct	cgt	tca	acg	gag	gtg	tat	gtc	cgg	tct	ata	ttt	ggc	917
Val	Pro	Tyr	Pro	Arg	Ser	Thr	Glu	Val	Tyr	Val	Arg	Ser	Ile	Phe	Gly	
	280					285					290					
cct	ggc	ggc	tcc	cca	tcc	ttg	ttt	cca	tca	gcc	tgc	tct	act	aaa	tct	965
Pro	Gly	Gly	Ser	Pro	Ser	Leu	Phe	Pro	Ser	Ala	Cys	Ser	Thr	Lys	Ser	
295				300						305					310	
acc	ttt	cat	gct	gtc	ccg	gtt	cac	atc	tgg	gat	crg	ctc	atg	ctc	ttt	1013
Thr	Phe	His	Ala	Val	Pro	Val	His	Ile	Trp	Asp	Xaa	Leu	Met	Leu	Phe	
			315						320					325		
ggg	gcc	acc	ctg	rac	gat	cag	gcg	ttc	tgc	tgt	tca	cgg	ctt	atg	act	1061
Gly	Ala	Thr	Leu	Xaa	Asp	Gln	Ala	Phe	Cys	Cys	Ser	Arg	Leu	Met	Thr	
			330				335						340			
tac	ctc	cgt	ggg	att	agt	tat	aag	gtc	act	gtc	ggg	gcg	ctt	gtc	gct	1109
Tyr	Leu	Arg	Gly	Ile	Ser	Tyr	Lys	Val	Thr	Val	Gly	Ala	Leu	Val	Ala	
		345					350					355				
aat	gag	ggg	tgg	aac	gcc	tct	gag	gat	gct	ctt	act	gca	gtg	atc	act	1157
Asn	Glu	Gly	Trp	Asn	Ala	Ser	Glu	Asp	Ala	Leu	Thr	Ala	Val	Ile	Thr	

360	365	370	
gcg gcc tat ctg acc atc tgc cat cag cgt tac ctt cgc acc cag gcg Ala Ala Tyr Leu Thr Ile Cys His Gln Arg Tyr Leu Arg Thr Gln Ala 375 380 385 390			1205
att tcc aag ggc atg cgc cgg ttg gag gtt gag cat gct cag aaa ttt Ile Ser Lys Gly Met Arg Arg Leu Glu Val Glu His Ala Gln Lys Phe 395 400 405			1253
atc aca aga ctc tac agc tgg cta ttt gag aag tct ggc cgt gac tac Ile Thr Arg Leu Tyr Ser Trp Leu Phe Glu Lys Ser Gly Arg Asp Tyr 410 415 420			1301
atc ccc ggc cgc cag ctt caa ttt tat gca caa tgc cga cgg tgg ctt Ile Pro Gly Arg Gln Leu Gln Phe Tyr Ala Gln Cys Arg Arg Trp Leu 425 430 435			1349
tct gca ggc ttc cac cta rac ccc agg rtg ctt gtc ttt gat gaa tca Ser Ala Gly Phe His Leu Xaa Pro Arg Xaa Leu Val Phe Asp Glu Ser 440 445 450			1397
gtg cca tgc cgt tgc agg acg ttt ttg aag aag gtc gcg ggt aaa ttc Val Pro Cys Arg Cys Arg Thr Phe Leu Lys Lys Val Ala Gly Lys Phe 455 460 465 470			1445
tgc tgt ttt atg cgg tgg ctg ggg cag gag tgt acc tgc ttc ttg gag Cys Cys Phe Met Arg Trp Leu Gly Gln Glu Cys Thr Cys Phe Leu Glu 475 480 485			1493
cca gcc gag ggt tta gtt ggt gat caa ggt cat gac aac gag gcc tat Pro Ala Glu Gly Leu Val Gly Asp Gln Gly His Asp Asn Glu Ala Tyr 490 495 500			1541
gaa ggt tct gag gtc gac cca gct gag cct gca cat ctt gat gtc tcg Glu Gly Ser Glu Val Asp Pro Ala Glu Pro Ala His Leu Asp Val Ser 505 510 515			1589
ggg act tat gcc gtc cat ggg cac cag ctt gag gcc ctc tat agg gca Gly Thr Tyr Ala Val His Gly His Gln Leu Glu Ala Leu Tyr Arg Ala 520 525 530			1637
ctt aat gtc cca cat gat att gcc gct cga gcc tcc cga cta acg gct Leu Asn Val Pro His Asp Ile Ala Ala Arg Ala Ser Arg Leu Thr Ala 535 540 545 550			1685
act gtt gag ctc gtt gct agt ccg gac cgc tta gag tgc cgc act gta Thr Val Glu Leu Val Ala Ser Pro Asp Arg Leu Glu Cys Arg Thr Val 555 560 565			1733
ctt ggt aat aag acc ttc cgg acg acg gtg gtt gat ggc gcc cat ctt Leu Gly Asn Lys Thr Phe Arg Thr Thr Val Val Asp Gly Ala His Leu 570 575 580			1781
gaa gcg aat ggc cct gag gag tat gtt ctg tca ttt gac gcc tct cgc Glu Ala Asn Gly Pro Glu Glu Tyr Val Leu Ser Phe Asp Ala Ser Arg 585 590 595			1829

cag tct atg ggg gcc ggg tcg cac agc ctc act tat gag ctc acc cct Gln Ser Met Gly Ala Gly Ser His Ser Leu Thr Tyr Glu Leu Thr Pro 600 605 610	1877
gcc ggt ctg cag gta aag att tca tct aat ggt ctg gat tgc act gcc Ala Gly Leu Gln Val Lys Ile Ser Ser Asn Gly Leu Asp Cys Thr Ala 615 620 625 630	1925
aca ttc ccc ccg ggt ggc gcc cct agc gcc gcg ccg ggg gag gtg gcs Thr Phe Pro Xaa Gly Ala Pro Ser Ala Ala Pro Gly Glu Val Xaa 635 640 645	1973
gcc ttc tgc agt gct ctt tat aga tac aat agg ttc acc cag cgg cat Ala Phe Cys Ser Ala Leu Tyr Arg Tyr Asn Arg Phe Thr Gln Arg His 650 655 660	2021
tcg ctg aca ggc gga cta tgg cta cat cct gag ggg ctg ctg ggt atc Ser Leu Thr Gly Gly Leu Trp Leu His Pro Glu Gly Leu Leu Gly Ile 665 670 675	2069
ttc ccc cca ttc tcc cct ggg cat att tgg gag tct gct aac ccc ttt Phe Pro Pro Phe Ser Pro Gly His Ile Trp Glu Ser Ala Asn Pro Phe 680 685 690	2117
tgc ggt gag ggg act ttg tat acc cga acc tgg tca acc tct ggt ttt Cys Gly Glu Gly Thr Leu Tyr Thr Arg Thr Trp Ser Thr Ser Gly Phe 695 700 705 710	2165
tct agt gat ttc tcc ccc cct gag gcg gcc gct cct gct tcg gct gcc Ser Ser Asp Phe Ser Pro Pro Glu Ala Ala Ala Pro Ala Ser Ala Ala 715 720 725	2213
gcc ccg ggg ttg ccc tac cct act cca cct gtt agt gat atc tgg gtg Ala Pro Gly Leu Pro Tyr Pro Thr Pro Pro Val Ser Asp Ile Trp Val 730 735 740	2261
tta cca ccg ccc tca gag gaa tct cat gtt gat gcg gca tct gta ccc Leu Pro Pro Pro Ser Glu Glu Ser His Val Asp Ala Ala Ser Val Pro 745 750 755	2309
tct gtt cct gag cct gct gga ttg acc agc cct att gtg ctt acc ccc Ser Val Pro Glu Pro Ala Gly Leu Thr Ser Pro Ile Val Leu Thr Pro 760 765 770	2357
ccc ccc ccc cct cct ccc gtg cgt aag ccg gca aca tcc ccg cct ccc Pro Pro Pro Pro Pro Pro Val Arg Lys Pro Ala Thr Ser Pro Pro Pro 775 780 785 790	2405
cgc act cgc cgt ctc ctt tac acc tac ccc gac ggc gcc aag gtg tat Arg Thr Arg Arg Leu Leu Tyr Thr Tyr Pro Asp Gly Ala Lys Val Tyr 795 800 805	2453
gcg ggg tca ttg tkt gag tca gac tgt gat tgg tta gtc aat gcc tca Ala Gly Ser Leu Xaa Glu Ser Asp Cys Asp Trp Leu Val Asn Ala Ser 810 815 820	2501

aac cct ggc cat cgc ccc ggg ggt ggc ctc tgc cat gct ttt tat caa	2549
Asn Pro Gly His Arg Pro Gly Gly Gly Leu Cys His Ala Phe Tyr Gln	
825 830 835	
cgt ttc cca gaa gcg ttc tac tcg act gaa ttc atc atg cgc gag ggc	2597
Arg Phe Pro Glu Ala Phe Tyr Ser Thr Glu Phe Ile Met Arg Glu Gly	
840 845 850	
ctt gca gca tac act tta acc ccg cgc cct att atc cat gca gtg gct	2645
Leu Ala Ala Tyr Thr Leu Thr Pro Arg Pro Ile Ile His Ala Val Ala	
855 860 865 870	
ccc gac tat agg gtt gag caa aac ccg aag agg ctt gag gca gcg tac	2693
Pro Asp Tyr Arg Val Glu Gln Asn Pro Lys Arg Leu Glu Ala Ala Tyr	
875 880 885	
cgg gaa act tgc tcc cgt cgt ggc acc gct gcc tac ccg ctt ttg ggc	2741
Arg Glu Thr Cys Ser Arg Arg Gly Thr Ala Ala Tyr Pro Leu Leu Gly	
890 895 900	
tcg ggt ata tac cag gtc cct gtt agc ctc agt ttt gat gcc tgg gaa	2789
Ser Gly Ile Tyr Gln Val Pro Val Ser Leu Ser Phe Asp Ala Trp Glu	
905 910 915	
cgc aat cac cgc ccc ggc gat gag ctt tac ttg aca gag ccc gcc gca	2837
Arg Asn His Arg Pro Gly Asp Glu Leu Tyr Leu Thr Glu Pro Ala Ala	
920 925 930	
gcc tgg ttt gag gct aat aag ccg gcg cag ccg gcg ctt act ata act	2885
Ala Trp Phe Glu Ala Asn Lys Pro Ala Gln Pro Ala Leu Thr Ile Thr	
935 940 945 950	
gag gac acg gcc cgt acg gcc aac ctg gca tta gag att gat gcc gcc	2933
Glu Asp Thr Ala Arg Thr Ala Asn Leu Ala Leu Glu Ile Asp Ala Ala	
955 960 965	
aca gag gtt ggc cgt gct tgt gcc ggc tgc acc atc agc ccc ggg att	2981
Thr Glu Val Gly Arg Ala Cys Ala Gly Cys Thr Ile Ser Pro Gly Ile	
970 975 980	
gtg cac tat cag ttt acc gcc ggg gtc ccg ggc tca ggc aag tca agg	3029
Val His Tyr Gln Phe Thr Ala Gly Val Pro Gly Ser Gly Lys Ser Arg	
985 990 995	
tcc ata caa cag gga gat gtc gat gtg gtg gtt gtg ccc acc cgg gag	3077
Ser Ile Gln Gln Gly Asp Val Asp Val Val Val Val Pro Thr Arg Glu	
1000 1005 1010	
ctc cgt aac agc tgg cgt cgc cgg ggt ttt gcg gcc ttc aca cct cac	3125
Leu Arg Asn Ser Trp Arg Arg Arg Gly Phe Ala Ala Phe Thr Pro His	
1015 1020 1025 1030	
aca gcg gcc cgt gtt act atc ggc cgc cgc gtt gtg att gat gag gct	3173
Thr Ala Ala Arg Val Thr Ile Gly Arg Arg Val Val Ile Asp Glu Ala	
1035 1040 1045	
cca tct ctc cca ccg cac ctg ctg ctg tta cac atg cag cgg gcc tcc	3221

Pro	Ser	Leu	Pro	Pro	His	Leu	Leu	Leu	Leu	His	Met	Gln	Arg	Ala	Ser		
			1050						1055						1060		
tcg	gtc	cat	ctc	ctt	ggt	gat	cca	aac	cag	att	cct	gct	att	gat	ttt	3269	
Ser	Val	His	Leu	Leu	Gly	Asp	Pro	Asn	Gln	Ile	Pro	Ala	Ile	Asp	Phe		
			1065						1070					1075			
gag	cat	gcc	ggc	ctg	gtc	ccc	gcg	atc	cgc	ccc	gag	ctt	gcg	cca	acg	3317	
Glu	His	Ala	Gly	Leu	Val	Pro	Ala	Ile	Arg	Pro	Glu	Leu	Ala	Pro	Thr		
			1080					1085					1090				
agc	tgg	tgg	cac	gtt	aca	cac	cgt	tgc	ccg	gcc	gat	gtg	tgc	gag	ctc	3365	
Ser	Trp	Trp	His	Val	Thr	His	Arg	Cys	Pro	Ala	Asp	Val	Cys	Glu	Leu		
							1100				1105				1110		
ata	cgt	ggg	gcc	tac	ccc	aaa	att	cag	acc	acg	agc	cgt	gtg	cta	cgg	3413	
Ile	Arg	Gly	Ala	Tyr	Pro	Lys	Ile	Gln	Thr	Thr	Ser	Arg	Val	Leu	Arg		
							1115				1120				1125		
tcc	ctg	ttt	tgg	aac	gaa	ccg	gcc	atc	ggc	caa	aag	ttg	gtt	ttt	acg	3461	
Ser	Leu	Phe	Trp	Asn	Glu	Pro	Ala	Ile	Gly	Gln	Lys	Leu	Val	Phe	Thr		
							1130							1140			
cag	gct	gct	aag	gct	gcc	aac	cct	ggt	gcg	att	acg	gtt	cac	gaa	gct	3509	
Gln	Ala	Ala	Lys	Ala	Ala	Asn	Pro	Gly	Ala	Ile	Thr	Val	His	Glu	Ala		
							1145							1155			
cag	ggt	gct	act	ttc	acg	gag	acc	aca	att	ata	gcc	acg	gcc	gac	gct	3557	
Gln	Gly	Ala	Thr	Phe	Thr	Glu	Thr	Thr	Ile	Ile	Ala	Thr	Ala	Asp	Ala		
														1160			
agg	ggc	ctc	att	cag	tca	tcc	cgg	gcc	cat	gct	ata	gtc	gca	ctc	acc	3605	
Arg	Gly	Leu	Ile	Gln	Ser	Ser	Arg	Ala	His	Ala	Ile	Val	Ala	Leu	Thr		
							1175							1180			
cgc	cat	act	gag	aag	tgt	gtt	att	ttg	gat	gcc	ccc	ggc	ttg	ttg	cgc	3653	
Arg	His	Thr	Glu	Lys	Cys	Val	Ile	Leu	Asp	Ala	Pro	Gly	Leu	Leu	Arg		
							1195				1200				1205		
gag	gtc	ggc	att	tcg	gat	gtt	att	gtc	aat	aac	ttt	ttc	ctt	gcc	ggt	3701	
Glu	Val	Gly	Ile	Ser	Asp	Val	Ile	Val	Asn	Asn	Phe	Phe	Leu	Ala	Gly		
							1210							1215			
gga	gag	gtc	ggc	cat	cac	cgc	cct	tct	gtg	ata	cct	cgc	ggc	aat	cct	3749	
Gly	Glu	Val	Gly	His	His	Arg	Pro	Ser	Val	Ile	Pro	Arg	Gly	Asn	Pro		
							1225							1230			
gat	cag	aac	ctc	ggg	act	cta	cag	gcc	ttt	ccg	ccg	tca	tgt	cag	atc	3797	
Asp	Gln	Asn	Leu	Gly	Thr	Leu	Gln	Ala	Phe	Pro	Pro	Ser	Cys	Gln	Ile		
							1240							1245			
agt	gct	tac	cat	cag	ttg	gct	gag	gaa	cta	ggt	cat	cgc	ccg	gcc	cct	3845	
Ser	Ala	Tyr	His	Gln	Leu	Ala	Glu	Glu	Leu	Gly	His	Arg	Pro	Ala	Pro		
							1255							1260			
gtc	gcc	gcc	gtc	ttg	ccc	cct	tgc	cct	gag	ctt	gag	cag	ggc	ctg	ctc	3893	
Val	Ala	Ala	Val	Leu	Pro	Pro	Cys	Pro	Glu	Leu	Glu	Gln	Gly	Leu	Leu		

1275	1280	1285	
tat atg cca caa gaa ctt act gtg tcc gat agc gtg ctg gtt ttt gag Tyr Met Pro Gln Glu Leu Thr Val Ser Asp Ser Val Leu Val Phe Glu 1290 1295 1300			3941
ctt acg gat ata gtc cac tgc cgt atg gcc gcc cca agc cag cga aag Leu Thr Asp Ile Val His Cys Arg Met Ala Ala Pro Ser Gln Arg Lys 1305 1310 1315			3989
gct gtt ctc tca acg ctt gtg ggg agg tac gcc cgt agg act aaa tta Ala Val Leu Ser Thr Leu Val Gly Arg Tyr Gly Arg Arg Thr Lys Leu 1320 1325 1330			4037
tat gag gcg gcg cat tca gat gtc cgt gag tcc cta gcg agg ttt atc Tyr Glu Ala Ala His Ser Asp Val Arg Glu Ser Leu Ala Arg Phe Ile 1335 1340 1345 1350			4085
ccc acc atc ggg cct gtt cgg gct acc aca tgt gag ctg tac gag ctg Pro Thr Ile Gly Pro Val Arg Ala Thr Thr Cys Glu Leu Tyr Glu Leu 1355 1360 1365			4133
gtt gaa gcc atg gta gag aag ggt cag gac gga tct gcc gtc cta gag Val Glu Ala Met Val Glu Lys Gly Gln Asp Gly Ser Ala Val Leu Glu 1370 1375 1380			4181
ctc gac ctt tgc aat cgt gac gtc tcg cgc atc aca ttt ttc caa aag Leu Asp Leu Cys Asn Arg Asp Val Ser Arg Ile Thr Phe Phe Gln Lys 1385 1390 1395			4229
gat tgc aat aag ttt aca act ggt gag act atc gcc cat gcc aag gtt Asp Cys Asn Lys Phe Thr Thr Gly Glu Thr Ile Ala His Gly Lys Val 1400 1405 1410			4277
ggc cag gcc ata tcg gcc tgg agc aag acc ttc tgt gct ctg ttt gcc Gly Gln Gly Ile Ser Ala Trp Ser Lys Thr Phe Cys Ala Leu Phe Gly 1415 1420 1425 1430			4325
ccg tgg ttc cgc gcc att gaa aag gaa ata ttg gcc cta ctc ccg cct Pro Trp Phe Arg Ala Ile Glu Lys Glu Ile Leu Ala Leu Leu Pro Pro 1435 1440 1445			4373
aat atc ttt tat ggc gac gcc tat gag gag tca gtg ttt gct gcc gct Asn Ile Phe Tyr Gly Asp Ala Tyr Glu Glu Ser Val Phe Ala Ala Ala 1450 1455 1460			4421
gtg tcc ggg gca ggg tca tgt atg gta ttt gaa aat gac ttc tca gag Val Ser Gly Ala Gly Ser Cys Met Val Phe Glu Asn Asp Phe Ser Glu 1465 1470 1475			4469
ttt gac agt acc cag aat aat ttc tct ctc gcc ctt gag tgt gtg gtt Phe Asp Ser Thr Gln Asn Asn Phe Ser Leu Gly Leu Glu Cys Val Val 1480 1485 1490			4517
atg gag gag tgc ggc atg ccc caa tgg tta att agg ttg tac cat ctg Met Glu Glu Cys Gly Met Pro Gln Trp Leu Ile Arg Leu Tyr His Leu 1495 1500 1505 1510			4565

gtc cgg tca gcc tgg att ttg cag gcg ccg aag gag tct ctt aag ggg	4613
Val Arg Ser Ala Trp Ile Leu Gln Ala Pro Lys Glu Ser Leu Lys Gly	
1515 1520 1525	
ttt tgg aag aag cac tct ggt gag cct ggt acc ctt ctc tgg aac act	4661
Phe Trp Lys Lys His Ser Gly Glu Pro Gly Thr Leu Leu Trp Asn Thr	
1530 1535 1540	
gtc tgg aac atg gcg att ata gca cat tgc tay gag ttc cgt gac ttt	4709
Val Trp Asn Met Ala Ile Ile Ala His Cys Xaa Glu Phe Arg Asp Phe	
1545 1550 1555	
cgt gtt gcc gcc ttc aag ggt gat gat tca gtg gtc ctc tgt agt gac	4757
Arg Val Ala Ala Phe Lys Gly Asp Asp Ser Val Val Leu Cys Ser Asp	
1560 1565 1570	
tac cga cag rgc cgt aac gcg gct gcc tta att gca ggc tgt ggg ctc	4805
Tyr Arg Gln Xaa Arg Asn Ala Ala Ala Leu Ile Ala Gly Cys Gly Leu	
1575 1580 1585 1590	
aaa ttg aag gtt gat tac cgc cct atc ggg cta tat gct gga gtg gtg	4853
Lys Leu Lys Val Asp Tyr Arg Pro Ile Gly Leu Tyr Ala Gly Val Val	
1595 1600 1605	
gtg gcc ccc ggt ttg ggg aca ctg ccc gat gtg gtg cgt ttt gcc ggt	4901
Val Ala Pro Gly Leu Gly Thr Leu Pro Asp Val Val Arg Phe Ala Gly	
1610 1615 1620	
cgg tta tct gag aag aat tgg ggc cct ggc ccg gag cgt gct gag cag	4949
Arg Leu Ser Glu Lys Asn Trp Gly Pro Gly Pro Glu Arg Ala Glu Gln	
1625 1630 1635	
ctg cgt ctt gct gtt tgt gat ttc ctt cga ggg ttg acg aat gtt gcg	4997
Leu Arg Leu Ala Val Cys Asp Phe Leu Arg Gly Leu Thr Asn Val Ala	
1640 1645 1650	
cag gtc tgt gtt gat gtt gtg tcc cgt gtc tat gga gtt agc ccc ggg	5045
Gln Val Cys Val Asp Val Val Ser Arg Val Tyr Gly Val Ser Pro Gly	
1655 1660 1665 1670	
ctg gta cat aac ctt att ggc atg ctg cag acc att gct gat ggc aag	5093
Leu Val His Asn Leu Ile Gly Met Leu Gln Thr Ile Ala Asp Gly Lys	
1675 1680 1685	
gcc cac ttt aca gar aat att aaa cct gtg ctt gac ctt aca aat tcc	5141
Ala His Phe Thr Xaa Asn Ile Lys Pro Val Leu Asp Leu Thr Asn Ser	
1690 1695 1700	
atc ata caa cgg gtg gaa tgaataacat gtcttttgca tcgcccattgg	5189
Ile Ile Gln Arg Val Glu	
1705	
gatcacc atg cgc cct agg gct gtt ctg ttg ttg ctc ttc gtg ctt ttg	5238
Met Arg Pro Arg Ala Val Leu Leu Leu Leu Phe Val Leu Leu	
1710 1715 1720	

cct atg ctg ccc gcg cca ccg gcc ggc cag ccg tct ggc cgc cgt cgt	5286
Pro Met Leu Pro Ala Pro Pro Ala Gly Gln Pro Ser Gly Arg Arg Arg	
1725 1730 1735	
ggg cgg cgc agc ggc ggt gcc ggc ggt ggt ttc tgg ggt gac agg gtt	5334
Gly Arg Arg Ser Gly Gly Ala Gly Gly Gly Phe Trp Gly Asp Arg Val	
1740 1745 1750	
gat tct cag ccc ttc gcc ctc ccc tat att cat cca acc aac ccc ttc	5382
Asp Ser Gln Pro Phe Ala Leu Pro Tyr Ile His Pro Thr Asn Pro Phe	
1755 1760 1765 1770	
gcc gcc gat gtc gtt tca caa ccc ggg gct gga act cgc cct cga cag	5430
Ala Ala Asp Val Val Ser Gln Pro Gly Ala Gly Thr Arg Pro Arg Gln	
1775 1780 1785	
ccg ccc cgc ccc ctt ggy tcc gct tgg cgt gac cag tcc cag cgc ccc	5478
Pro Pro Arg Pro Leu Xaa Ser Ala Trp Arg Asp Gln Ser Gln Arg Pro	
1790 1795 1800	
tcc gct gcc ccc cgt cgt cga tct gcc cca gct ggg gct gcg ccg ctg	5526
Ser Ala Ala Pro Arg Arg Arg Ser Ala Pro Ala Gly Ala Ala Pro Leu	
1805 1810 1815	
act gcc gtg tca ccg gct cct gac aca gcc cct gta cct gat gtt gac	5574
Thr Ala Val Ser Pro Ala Pro Asp Thr Ala Pro Val Pro Asp Val Asp	
1820 1825 1830	
tca cgt ggt gct att ctg cgc cgg cag tac aat ttg tcc acg tcc ccg	5622
Ser Arg Gly Ala Ile Leu Arg Arg Gln Tyr Asn Leu Ser Thr Ser Pro	
1835 1840 1845 1850	
ctc acg tca tct gtc gct tgc ggt act aat ttg gtc ctc tat gct gcc	5670
Leu Thr Ser Ser Val Ala Ser Gly Thr Asn Leu Val Leu Tyr Ala Ala	
1855 1860 1865	
ccg ctg aat ccc ctc ttg cct ctc cag gat ggt acc aac act cat att	5718
Pro Leu Asn Pro Leu Leu Pro Leu Gln Asp Gly Thr Asn Thr His Ile	
1870 1875 1880	
atg gct act gag gca tcc aat tat gcc cag tat cgg gtt gtt cga gct	5766
Met Ala Thr Glu Ala Ser Asn Tyr Ala Gln Tyr Arg Val Val Arg Ala	
1885 1890 1895	
aca atc cgt tat cgc ccg ctg gtg ccg aat gcc gtt ggt ggc tat gcc	5814
Thr Ile Arg Tyr Arg Pro Leu Val Pro Asn Ala Val Gly Gly Tyr Ala	
1900 1905 1910	
att tcc att tct ttc tgg ccc caa act aca act acc cct act tct gtc	5862
Ile Ser Ile Ser Phe Trp Pro Gln Thr Thr Thr Thr Pro Thr Ser Val	
1915 1920 1925 1930	
gat atg aat tct att act tcc acy gat gtt agg att ttg gtt cag ccc	5910
Asp Met Asn Ser Ile Thr Ser Xaa Asp Val Arg Ile Leu Val Gln Pro	
1935 1940 1945	
ggg att gcc tcc gag cta gtc atc ccc agt gag cgc ctt cat tac cgt	5958

Gly	Ile	Ala	Ser	Glu	Leu	Val	Ile	Pro	Ser	Glu	Arg	Leu	His	Tyr	Arg	
			1950					1955					1960			
aat	caa	ggc	tgg	cgc	tct	gtt	gag	acc	acg	ggg	gtg	gct	gag	gag	gag	6006
Asn	Gln	Gly	Trp	Arg	Ser	Val	Glu	Thr	Thr	Gly	Val	Ala	Glu	Glu	Glu	
		1965					1970					1975				
gct	act	tcc	ggg	ctg	gta	atg	ctt	tgc	att	cat	ggc	tct	cct	gtt	aat	6054
Ala	Thr	Ser	Gly	Leu	Val	Met	Leu	Cys	Ile	His	Gly	Ser	Pro	Val	Asn	
		1980				1985					1990					
tcc	tac	act	aat	aca	cct	tac	act	ggg	gag	ctg	ggg	ctt	ctt	gat	ttt	6102
Ser	Tyr	Thr	Asn	Thr	Pro	Tyr	Thr	Gly	Ala	Leu	Gly	Leu	Leu	Asp	Phe	
1995					2000					2005					2010	
gca	cta	gag	ctt	gaa	ttt	agg	aat	ttg	aca	ccc	ggg	aac	acc	aac	acc	6150
Ala	Leu	Glu	Leu	Glu	Phe	Arg	Asn	Leu	Thr	Pro	Gly	Asn	Thr	Asn	Thr	
				2015					2020					2025		
cgt	gtt	tcc	cgg	tat	acc	agc	aca	gcc	cgc	cac	cgg	ctg	cgc	cgt	ggg	6198
Arg	Val	Ser	Arg	Tyr	Thr	Ser	Thr	Ala	Arg	His	Arg	Leu	Arg	Arg	Gly	
			2030					2035					2040			
gct	gat	ggg	act	gct	gag	ctt	act	acc	aca	gca	gcc	aca	cgt	ttc	atg	6246
Ala	Asp		Thr	Ala	Glu	Leu	Thr	Thr	Thr	Ala	Ala	Thr	Arg	Phe	Met	
		2045					2050					2055				
aag	gac	ctg	cac	ttc	gct	ggc	acg	aat	ggc	gtt	ggg	gag	gtg	ggg	cgt	6294
Lys	Asp	Leu	His	Phe	Ala	Gly	Thr	Asn	Gly	Val	Gly	Glu	Val	Gly	Arg	
	2060					2065					2070					
ggg	atc	gcc	ctg	aca	ctg	ttc	aat	ctc	gct	gat	acg	ctt	ctc	ggc	ggg	6342
Gly	Ile	Ala	Leu	Thr	Leu	Phe	Asn	Leu	Ala	Asp	Thr	Leu	Leu	Gly	Gly	
2075					2080					2085					2090	
tta	ccg	aca	gaa	ttg	att	tgc	tgc	gct	ggg	ggc	caa	ctg	ttt	tac	tcc	6390
Leu	Pro	Thr	Glu	Leu	Ile	Ser	Ser	Ala	Gly	Gly	Gln	Leu	Phe	Tyr	Ser	
				2095					2100					2105		
cgc	ccg	gtt	gtc	tca	gcc	aat	ggc	gag	cca	aca	gta	aag	tta	tat	aca	6438
Arg	Pro	Val	Val	Ser	Ala	Asn	Gly	Glu	Pro	Thr	Val	Lys	Leu	Tyr	Thr	
			2110					2115					2120			
tct	gtt	gag	aat	gag	cag	caa	gac	aag	ggc	atc	acc	att	cca	cat	gat	6486
Ser	Val	Glu	Asn	Ala	Gln	Gln	Asp	Lys	Gly	Ile	Thr	Ile	Pro	His	Asp	
		2125					2130					2135				
ata	gac	ctg	ggg	gac	tcc	cgt	gtg	gtt	atc	cag	gat	tat	gat	aac	cag	6534
Ile	Asp	Leu	Gly	Asp	Ser	Arg	Val	Val	Ile	Gln	Asp	Tyr	Asp	Asn	Gln	
		2140				2145					2150					
cay	gag	caa	gac	cga	cct	act	ccg	tca	cct	gcc	ccc	tct	cgc	ccc	ttc	6582
Xaa	Glu	Gln	Asp	Arg	Pro	Thr	Pro	Ser	Pro	Ala	Pro	Ser	Arg	Pro	Phe	
2155					2160					2165					2170	
tca	gtt	ctt	cgt	gcc	aat	gat	gtt	ttg	tgg	ctt	tcc	ctc	act	gcc	gct	6630
Ser	Val	Leu	Arg	Ala	Asn	Asp	Val	Leu	Trp	Leu	Ser	Leu	Thr	Ala	Ala	

2175	2180	2185	
gag tat gac cag act acg tat ggg tgc tcc acc aac cct atg tat gtc Glu Tyr Asp Gln Thr Thr Tyr Gly Ser Ser Thr Asn Pro Met Tyr Val 2190 2195 2200			6678
tct gac aca gtt acg ctt gtt aat gtg gct act ggt gct cag gct gtt Ser Asp Thr Val Thr Leu Val Asn Val Ala Thr Gly Ala Gln Ala Val 2205 2210 2215			6726
gcc cgc tcc ctt gat tgg tct aaa gtt act ctg gac ggc cgc ccc ctt Ala Arg Ser Leu Asp Trp Ser Lys Val Thr Leu Asp Gly Arg Pro Leu 2220 2225 2230			6774
act acc att cag cag tat tct aag aca ttt tat gtt ctg ccg ctg cgc Thr Thr Ile Gln Gln Tyr Ser Lys Thr Phe Tyr Val Leu Pro Leu Arg 2235 2240 2245 2250			6822
ggg aag ctg tcc ttt tgg gag gct ggc acg act aag gcc ggc tac cct Gly Lys Leu Ser Phe Trp Glu Ala Gly Thr Thr Lys Ala Gly Tyr Pro 2255 2260 2265			6870
tac aat tat aat act acc gct agt gac caa att ttg att gag aat gcg Tyr Asn Tyr Asn Thr Thr Ala Ser Asp Gln Ile Leu Ile Glu Asn Ala 2270 2275 2280			6918
gcc ggc cac cgt gtc gct att tcc acc tat acc act agc tta ggt gcc Ala Gly His Arg Val Ala Ile Ser Thr Tyr Thr Thr Ser Leu Gly Ala 2285 2290 2295			6966
ggt cct acc tgc atc tct gcg gtc ggc gta ctg gct cca cac tct gcc Gly Pro Thr Ser Ile Ser Ala Val Gly Val Leu Ala Pro His Ser Ala 2300 2305 2310			7014
ctt gcc gtt ctt gag gat act att gat tac ccc gcc cgt gcc cat act Leu Ala Val Leu Glu Asp Thr Ile Asp Tyr Pro Ala Arg Ala His Thr 2315 2320 2325 2330			7062
ttt gat gat ttt tgc ccg gag tgc cgt acc cta ggt ttg cag ggt tgt Phe Asp Asp Phe Cys Pro Glu Cys Arg Thr Leu Gly Leu Gln Gly Cys 2335 2340 2345			7110
gca ttc cag tct act att gct gag ctg cag cgt tta aaa atg aag gta Ala Phe Gln Ser Thr Ile Ala Glu Leu Gln Arg Leu Lys Met Lys Val 2350 2355 2360			7158
ggt aaa acc cgg gag tct taattaattc cttctgtgcc cccttcgtag Gly Lys Thr Arg Glu Ser 2365			7206
tttctttcgc ttttatttct tattttctgct ttccgcgctc cctggaaaaa aaaaaaaaaa aaaaaaaaaa a			7266 7277

<210> 166

<211> 1708

<212> PRT

<213> Hepatitis E Virus

<220>
 <223> Xaa = Unknown or Other at position 322
 <223> Xaa = Unknown or Other at position 331
 <223> Xaa = Unknown or Other at position 445
 <223> Xaa = Unknown or Other at position 448
 <223> Xaa = Unknown or Other at position 634
 <223> Xaa = Unknown or Other at position 646
 <223> Xaa = Unknown or Other at position 811
 <223> Xaa = Unknown or Other at position 1553
 <223> Xaa = Unknown or Other at position 1578
 <223> Xaa = Unknown or Other at position 1691

<400> 166
 Met Glu Ala His Gln Phe Ile Lys Ala Pro Gly Ile Thr Thr Ala Ile
 1 5 10 15
 Glu Gln Ala Ala Leu Ala Ala Ala Asn Ser Ala Leu Ala Asn Ala Val
 20 25 30
 Val Val Arg Pro Phe Leu Ser Arg Val Gln Thr Glu Ile Leu Ile Asn
 35 40 45
 Leu Met Gln Pro Arg Gln Leu Val Phe Arg Pro Glu Val Leu Trp Asn
 50 55 60
 His Pro Ile Gln Arg Val Ile His Asn Glu Leu Glu Gln Tyr Cys Arg
 65 70 75 80
 Ala Arg Ala Gly Arg Cys Leu Glu Val Gly Ala His Pro Arg Ser Ile
 85 90 95
 Asn Asp Asn Pro Asn Val Leu His Arg Cys Phe Leu Arg Pro Val Gly
 100 105 110
 Arg Asp Val Gln Arg Trp Tyr Ser Ala Pro Thr Arg Gly Pro Ala Ala
 115 120 125
 Asn Cys Arg Arg Ser Ala Leu Arg Gly Leu Pro Pro Val Asp Arg Thr
 130 135 140
 Tyr Cys Phe Asp Gly Phe Ser Arg Cys Ala Phe Ala Ala Glu Thr Gly
 145 150 155 160
 Val Ala Leu Tyr Ser Leu His Asp Leu Trp Pro Ala Asp Val Ala Glu
 165 170 175
 Ala Met Ala Arg His Gly Met Thr Arg Leu Tyr Ala Ala Leu His Leu
 180 185 190
 Pro Pro Glu Val Leu Leu Pro Pro Gly Thr Tyr His Thr Thr Ser Tyr
 195 200 205
 Leu Leu Ile His Asp Gly Asn Arg Ala Val Val Thr Tyr Glu Gly Asp
 210 215 220
 Thr Ser Ala Gly Tyr Asn His Asp Val Ser Ile Leu Arg Ala Trp Ile
 225 230 235 240
 Arg Thr Thr Lys Ile Val Gly Asp His Pro Leu Val Ile Glu Arg Val
 245 250 255
 Arg Ala Ile Gly Cys His Phe Val Leu Leu Leu Thr Ala Ala Pro Glu
 260 265 270

Pro	Ser	Pro	Met	Pro	Tyr	Val	Pro	Tyr	Pro	Arg	Ser	Thr	Glu	Val	Tyr
		275					280					285			
Val	Arg	Ser	Ile	Phe	Gly	Pro	Gly	Gly	Ser	Pro	Ser	Leu	Phe	Pro	Ser
	290					295				300					
Ala	Cys	Ser	Thr	Lys	Ser	Thr	Phe	His	Ala	Val	Pro	Val	His	Ile	Trp
305					310					315					320
Asp	Xaa	Leu	Met	Leu	Phe	Gly	Ala	Thr	Leu	Xaa	Asp	Gln	Ala	Phe	Cys
			325						330					335	
Cys	Ser	Arg	Leu	Met	Thr	Tyr	Leu	Arg	Gly	Ile	Ser	Tyr	Lys	Val	Thr
		340						345					350		
Val	Gly	Ala	Leu	Val	Ala	Asn	Glu	Gly	Trp	Asn	Ala	Ser	Glu	Asp	Ala
	355						360					365			
Leu	Thr	Ala	Val	Ile	Thr	Ala	Ala	Tyr	Leu	Thr	Ile	Cys	His	Gln	Arg
	370					375					380				
Tyr	Leu	Arg	Thr	Gln	Ala	Ile	Ser	Lys	Gly	Met	Arg	Arg	Leu	Glu	Val
385				390						395					400
Glu	His	Ala	Gln	Lys	Phe	Ile	Thr	Arg	Leu	Tyr	Ser	Trp	Leu	Phe	Glu
			405					410						415	
Lys	Ser	Gly	Arg	Asp	Tyr	Ile	Pro	Gly	Arg	Gln	Leu	Gln	Phe	Tyr	Ala
		420					425						430		
Gln	Cys	Arg	Arg	Trp	Leu	Ser	Ala	Gly	Phe	His	Leu	Xaa	Pro	Arg	Xaa
	435						440					445			
Leu	Val	Phe	Asp	Glu	Ser	Val	Pro	Cys	Arg	Cys	Arg	Thr	Phe	Leu	Lys
	450					455					460				
Lys	Val	Ala	Gly	Lys	Phe	Cys	Cys	Phe	Met	Arg	Trp	Leu	Gly	Gln	Glu
465				470						475					480
Cys	Thr	Cys	Phe	Leu	Glu	Pro	Ala	Glu	Gly	Leu	Val	Gly	Asp	Gln	Gly
			485					490						495	
His	Asp	Asn	Glu	Ala	Tyr	Glu	Gly	Ser	Glu	Val	Asp	Pro	Ala	Glu	Pro
		500						505					510		
Ala	His	Leu	Asp	Val	Ser	Gly	Thr	Tyr	Ala	Val	His	Gly	His	Gln	Leu
	515						520					525			
Glu	Ala	Leu	Tyr	Arg	Ala	Leu	Asn	Val	Pro	His	Asp	Ile	Ala	Ala	Arg
	530					535					540				
Ala	Ser	Arg	Leu	Thr	Ala	Thr	Val	Glu	Leu	Val	Ala	Ser	Pro	Asp	Arg
545				550						555					560
Leu	Glu	Cys	Arg	Thr	Val	Leu	Gly	Asn	Lys	Thr	Phe	Arg	Thr	Thr	Val
			565						570					575	
Val	Asp	Gly	Ala	His	Leu	Glu	Ala	Asn	Gly	Pro	Glu	Glu	Tyr	Val	Leu
		580						585					590		
Ser	Phe	Asp	Ala	Ser	Arg	Gln	Ser	Met	Gly	Ala	Gly	Ser	His	Ser	Leu
	595						600					605			
Thr	Tyr	Glu	Leu	Thr	Pro	Ala	Gly	Leu	Gln	Val	Lys	Ile	Ser	Ser	Asn
	610					615						620			
Gly	Leu	Asp	Cys	Thr	Ala	Thr	Phe	Pro	Xaa	Gly	Gly	Ala	Pro	Ser	Ala
625				630						635					640
Ala	Pro	Gly	Glu	Val	Xaa	Ala	Phe	Cys	Ser	Ala	Leu	Tyr	Arg	Tyr	Asn
			645						650					655	
Arg	Phe	Thr	Gln	Arg	His	Ser	Leu	Thr	Gly	Gly	Leu	Trp	Leu	His	Pro
		660						665					670		
Glu	Gly	Leu	Leu	Gly	Ile	Phe	Pro	Pro	Phe	Ser	Pro	Gly	His	Ile	Trp
	675						680					685			
Glu	Ser	Ala	Asn	Pro	Phe	Cys	Gly	Glu	Gly	Thr	Leu	Tyr	Thr	Arg	Thr
	690					695					700				
Trp	Ser	Thr	Ser	Gly	Phe	Ser	Ser	Asp	Phe	Ser	Pro	Pro	Glu	Ala	Ala
705				710						715					720
Ala	Pro	Ala	Ser	Ala	Ala	Ala	Pro	Gly	Leu	Pro	Tyr	Pro	Thr	Pro	Pro

Val	Ser	Asp	Ile	Trp	Val	Leu	Pro	Pro	Pro	Ser	Glu	Glu	Ser	His	Val
			740					745					750		
Asp	Ala	Ala	Ser	Val	Pro	Ser	Val	Pro	Glu	Pro	Ala	Gly	Leu	Thr	Ser
		755					760					765			
Pro	Ile	Val	Leu	Thr	Pro	Pro	Pro	Pro	Pro	Pro	Pro	Val	Arg	Lys	Pro
		770					775					780			
Ala	Thr	Ser	Pro	Pro	Pro	Arg	Thr	Arg	Arg	Leu	Leu	Tyr	Thr	Tyr	Pro
785						790				795					800
Asp	Gly	Ala	Lys	Val	Tyr	Ala	Gly	Ser	Leu	Xaa	Glu	Ser	Asp	Cys	Asp
			805						810					815	
Trp	Leu	Val	Asn	Ala	Ser	Asn	Pro	Gly	His	Arg	Pro	Gly	Gly	Gly	Leu
			820					825					830		
Cys	His	Ala	Phe	Tyr	Gln	Arg	Phe	Pro	Glu	Ala	Phe	Tyr	Ser	Thr	Glu
		835					840					845			
Phe	Ile	Met	Arg	Glu	Gly	Leu	Ala	Ala	Tyr	Thr	Leu	Thr	Pro	Arg	Pro
		850				855					860				
Ile	Ile	His	Ala	Val	Ala	Pro	Asp	Tyr	Arg	Val	Glu	Gln	Asn	Pro	Lys
865					870					875					880
Arg	Leu	Glu	Ala	Ala	Tyr	Arg	Glu	Thr	Cys	Ser	Arg	Arg	Gly	Thr	Ala
			885						890					895	
Ala	Tyr	Pro	Leu	Leu	Gly	Ser	Gly	Ile	Tyr	Gln	Val	Pro	Val	Ser	Leu
			900					905					910		
Ser	Phe	Asp	Ala	Trp	Glu	Arg	Asn	His	Arg	Pro	Gly	Asp	Glu	Leu	Tyr
		915					920					925			
Leu	Thr	Glu	Pro	Ala	Ala	Ala	Trp	Phe	Glu	Ala	Asn	Lys	Pro	Ala	Gln
		930				935					940				
Pro	Ala	Leu	Thr	Ile	Thr	Glu	Asp	Thr	Ala	Arg	Thr	Ala	Asn	Leu	Ala
945					950					955					960
Leu	Glu	Ile	Asp	Ala	Ala	Thr	Glu	Val	Gly	Arg	Ala	Cys	Ala	Gly	Cys
			965						970					975	
Thr	Ile	Ser	Pro	Gly	Ile	Val	His	Tyr	Gln	Phe	Thr	Ala	Gly	Val	Pro
			980					985					990		
Gly	Ser	Gly	Lys	Ser	Arg	Ser	Ile	Gln	Gln	Gly	Asp	Val	Asp	Val	Val
		995					1000					1005			
Val	Val	Pro	Thr	Arg	Glu	Leu	Arg	Asn	Ser	Trp	Arg	Arg	Arg	Gly	Phe
		1010				1015					1020				
Ala	Ala	Phe	Thr	Pro	His	Thr	Ala	Ala	Arg	Val	Thr	Ile	Gly	Arg	Arg
1025					1030					1035					1040
Val	Val	Ile	Asp	Glu	Ala	Pro	Ser	Leu	Pro	Pro	His	Leu	Leu	Leu	Leu
			1045						1050					1055	
His	Met	Gln	Arg	Ala	Ser	Ser	Val	His	Leu	Leu	Gly	Asp	Pro	Asn	Gln
			1060					1065					1070		
Ile	Pro	Ala													

Ala Ile Val Ala Leu Thr Arg His Thr Glu Lys Cys Val Ile Leu Asp
 1185 1190 1195 1200
 Ala Pro Gly Leu Leu Arg Glu Val Gly Ile Ser Asp Val Ile Val Asn
 1205 1210 1215
 Asn Phe Phe Leu Ala Gly Gly Glu Val Gly His His Arg Pro Ser Val
 1220 1225 1230
 Ile Pro Arg Gly Asn Pro Asp Gln Asn Leu Gly Thr Leu Gln Ala Phe
 1235 1240 1245
 Pro Pro Ser Cys Gln Ile Ser Ala Tyr His Gln Leu Ala Glu Glu Leu
 1250 1255 1260
 Gly His Arg Pro Ala Pro Val Ala Ala Val Leu Pro Pro Cys Pro Glu
 1265 1270 1275 1280
 Leu Glu Gln Gly Leu Leu Tyr Met Pro Gln Glu Leu Thr Val Ser Asp
 1285 1290 1295
 Ser Val Leu Val Phe Glu Leu Thr Asp Ile Val His Cys Arg Met Ala
 1300 1305 1310
 Ala Pro Ser Gln Arg Lys Ala Val Leu Ser Thr Leu Val Gly Arg Tyr
 1315 1320 1325
 Gly Arg Arg Thr Lys Leu Tyr Glu Ala Ala His Ser Asp Val Arg Glu
 1330 1335 1340
 Ser Leu Ala Arg Phe Ile Pro Thr Ile Gly Pro Val Arg Ala Thr Thr
 1345 1350 1355 1360
 Cys Glu Leu Tyr Glu Leu Val Glu Ala Met Val Glu Lys Gly Gln Asp
 1365 1370 1375
 Gly Ser Ala Val Leu Glu Leu Asp Leu Cys Asn Arg Asp Val Ser Arg
 1380 1385 1390
 Ile Thr Phe Phe Gln Lys Asp Cys Asn Lys Phe Thr Thr Gly Glu Thr
 1395 1400 1405
 Ile Ala His Gly Lys Val Gly Gln Gly Ile Ser Ala Trp Ser Lys Thr
 1410 1415 1420
 Phe Cys Ala Leu Phe Gly Pro Trp Phe Arg Ala Ile Glu Lys Glu Ile
 1425 1430 1435 1440
 Leu Ala Leu Leu Pro Asn Ile Phe Tyr Gly Asp Ala Tyr Glu Glu
 1445 1450 1455
 Ser Val Phe Ala Ala Ala Val Ser Gly Ala Gly Ser Cys Met Val Phe
 1460 1465 1470
 Glu Asn Asp Phe Ser Glu Phe Asp Ser Thr Gln Asn Asn Phe Ser Leu
 1475 1480 1485
 Gly Leu Glu Cys Val Val Met Glu Glu Cys Gly Met Pro Gln Trp Leu
 1490 1495 1500
 Ile Arg Leu Tyr His Leu Val Arg Ser Ala Trp Ile Leu Gln Ala Pro
 1505 1510 1515 1520
 Lys Glu Ser Leu Lys Gly Phe Trp Lys Lys His Ser Gly Glu Pro Gly
 1525 1530 1535
 Thr Leu Leu Trp Asn Thr Val Trp Asn Met Ala Ile Ile Ala His Cys
 1540 1545 1550
 Xaa Glu Phe Arg Asp Phe Arg Val Ala Ala Phe Lys Gly Asp Asp Ser
 1555 1560 1565
 Val Val Leu Cys Ser Asp Tyr Arg Gln Xaa Arg Asn Ala Ala Ala Leu
 1570 1575 1580
 Ile Ala Gly Cys Gly Leu Lys Leu Lys Val Asp Tyr Arg Pro Ile Gly
 1585 1590 1595 1600
 Leu Tyr Ala Gly Val Val Val Ala Pro Gly Leu Gly Thr Leu Pro Asp
 1605 1610 1615
 Val Val Arg Phe Ala Gly Arg Leu Ser Glu Lys Asn Trp Gly Pro Gly
 1620 1625 1630
 Pro Glu Arg Ala Glu Gln Leu Arg Leu Ala Val Cys Asp Phe Leu Arg

1635 1640 1645
 Gly Leu Thr Asn Val Ala Gln Val Cys Val Asp Val Val Ser Arg Val
 1650 1655 1660
 Tyr Gly Val Ser Pro Gly Leu Val His Asn Leu Ile Gly Met Leu Gln
 1665 1670 1675 1680
 Thr Ile Ala Asp Gly Lys Ala His Phe Thr Xaa Asn Ile Lys Pro Val
 1685 1690 1695
 Leu Asp Leu Thr Asn Ser Ile Ile Gln Arg Val Glu
 1700 1705

<210> 167
 <211> 660
 <212> PRT
 <213> Hepatitis E Virus

<220>
 <223> Xaa = Unknown or Other at position 84
 <223> Xaa = Unknown or Other at position 230
 <223> Xaa = Unknown or Other at position 447

<400> 167
 Met Arg Pro Arg Ala Val Leu Leu Leu Leu Phe Val Leu Leu Pro Met
 1 5 10 15
 Leu Pro Ala Pro Pro Ala Gly Gln Pro Ser Gly Arg Arg Arg Gly Arg
 20 25 30
 Arg Ser Gly Gly Ala Gly Gly Gly Phe Trp Gly Asp Arg Val Asp Ser
 35 40 45
 Gln Pro Phe Ala Leu Pro Tyr Ile His Pro Thr Asn Pro Phe Ala Ala
 50 55 60
 Asp Val Val Ser Gln Pro Gly Ala Gly Thr Arg Pro Arg Gln Pro Pro
 65 70 75 80
 Arg Pro Leu Xaa Ser Ala Trp Arg Asp Gln Ser Gln Arg Pro Ser Ala
 85 90 95
 Ala Pro Arg Arg Arg Ser Ala Pro Ala Gly Ala Ala Pro Leu Thr Ala
 100 105 110
 Val Ser Pro Ala Pro Asp Thr Ala Pro Val Pro Asp Val Asp Ser Arg
 115 120 125
 Gly Ala Ile Leu Arg Arg Gln Tyr Asn Leu Ser Thr Ser Pro Leu Thr
 130 135 140
 Ser Ser Val Ala Ser Gly Thr Asn Leu Val Leu Tyr Ala Ala Pro Leu
 145 150 155 160
 Asn Pro Leu Leu Pro Leu Gln Asp Gly Thr Asn Thr His Ile Met Ala
 165 170 175
 Thr Glu Ala Ser Asn Tyr Ala Gln Tyr Arg Val Val Arg Ala Thr Ile
 180 185 190
 Arg Tyr Arg Pro Leu Val Pro Asn Ala Val Gly Gly Tyr Ala Ile Ser
 195 200 205
 Ile Ser Phe Trp Pro Gln Thr Thr Thr Thr Pro Thr Ser Val Asp Met
 210 215 220
 Asn Ser Ile Thr Ser Xaa Asp Val Arg Ile Leu Val Gln Pro Gly Ile
 225 230 235 240
 Ala Ser Glu Leu Val Ile Pro Ser Glu Arg Leu His Tyr Arg Asn Gln
 245 250 255
 Gly Trp Arg Ser Val Glu Thr Thr Gly Val Ala Glu Glu Glu Ala Thr
 260 265 270

```

Ser Gly Leu Val Met Leu Cys Ile His Gly Ser Pro Val Asn Ser Tyr
    275                280                285
Thr Asn Thr Pro Tyr Thr Gly Ala Leu Gly Leu Leu Asp Phe Ala Leu
    290                295                300
Glu Leu Glu Phe Arg Asn Leu Thr Pro Gly Asn Thr Asn Thr Arg Val
305                310                315                320
Ser Arg Tyr Thr Ser Thr Ala Arg His Arg Leu Arg Arg Gly Ala Asp
    325                330                335
Gly Thr Ala Glu Leu Thr Thr Thr Ala Ala Thr Arg Phe Met Lys Asp
    340                345                350
Leu His Phe Ala Gly Thr Asn Gly Val Gly Glu Val Gly Arg Gly Ile
    355                360                365
Ala Leu Thr Leu Phe Asn Leu Ala Asp Thr Leu Leu Gly Gly Leu Pro
    370                375                380
Thr Glu Leu Ile Ser Ser Ala Gly Gly Gln Leu Phe Tyr Ser Arg Pro
385                390                395                400
Val Val Ser Ala Asn Gly Glu Pro Thr Val Lys Leu Tyr Thr Ser Val
    405                410                415
Glu Asn Ala Gln Gln Asp Lys Gly Ile Thr Ile Pro His Asp Ile Asp
    420                425                430
Leu Gly Asp Ser Arg Val Val Ile Gln Asp Tyr Asp Asn Gln Xaa Glu
    435                440                445
Gln Asp Arg Pro Thr Pro Ser Pro Ala Pro Ser Arg Pro Phe Ser Val
    450                455                460
Leu Arg Ala Asn Asp Val Leu Trp Leu Ser Leu Thr Ala Ala Glu Tyr
465                470                475                480
Asp Gln Thr Thr Tyr Gly Ser Ser Thr Asn Pro Met Tyr Val Ser Asp
    485                490                495
Thr Val Thr Leu Val Asn Val Ala Thr Gly Ala Gln Ala Val Ala Arg
    500                505                510
Ser Leu Asp Trp Ser Lys Val Thr Leu Asp Gly Arg Pro Leu Thr Thr
    515                520                525
Ile Gln Gln Tyr Ser Lys Thr Phe Tyr Val Leu Pro Leu Arg Gly Lys
    530                535                540
Leu Ser Phe Trp Glu Ala Gly Thr Thr Lys Ala Gly Tyr Pro Tyr Asn
545                550                555                560
Tyr Asn Thr Thr Ala Ser Asp Gln Ile Leu Ile Glu Asn Ala Ala Gly
    565                570                575
His Arg Val Ala Ile Ser Thr Tyr Thr Thr Ser Leu Gly Ala Gly Pro
    580                585                590
Thr Ser Ile Ser Ala Val Gly Val Leu Ala Pro His Ser Ala Leu Ala
    595                600                605
Val Leu Glu Asp Thr Ile Asp Tyr Pro Ala Arg Ala His Thr Phe Asp
    610                615                620
Asp Phe Cys Pro Glu Cys Arg Thr Leu Gly Leu Gln Gly Cys Ala Phe
625                630                635                640
Gln Ser Thr Ile Ala Glu Leu Gln Arg Leu Lys Met Lys Val Gly Lys
    645                650                655
Thr Arg Glu Ser
    660

```

<210> 168

<211> 122

<212> PRT

<213> Hepatitis E Virus

<220>

<223> us2 orf3

<223> Xaa = Unknown or Other at position 97

<400> 168

```

Met Asn Asn Met Ser Phe Ala Ser Pro Met Gly Ser Pro Cys Ala Leu
 1          5          10          15
Gly Leu Phe Cys Cys Cys Ser Ser Cys Phe Cys Leu Cys Cys Pro Arg
          20          25          30
His Arg Pro Ala Ser Arg Leu Ala Val Val Gly Gly Ala Ala Ala
          35          40          45
Val Pro Ala Val Val Ser Gly Val Thr Gly Leu Ile Leu Ser Pro Ser
          50          55          60
Pro Ser Pro Ile Phe Ile Gln Pro Thr Pro Ser Pro Pro Met Ser Phe
65          70          75          80
His Asn Pro Gly Leu Glu Leu Ala Leu Asp Ser Arg Pro Ala Pro Leu
          85          90          95
Xaa Pro Leu Gly Val Thr Ser Pro Ser Ala Pro Pro Leu Pro Pro Val
          100          105          110
Val Asp Leu Pro Gln Leu Gly Leu Arg Arg
          115          120

```

<210> 169

<211> 33

<212> PRT

<213> Hepatitis E Virus

<220>

<223> M 4-2

<400> 169

```

Ala Asn Gln Pro Gly His Leu Ala Pro Leu Gly Glu Ile Arg Pro Ser
 1          5          10          15
Ala Pro Pro Leu Pro Pro Val Ala Asp Leu Pro Gln Pro Gly Leu Arg
          20          25          30
Arg

```

<210> 170

<211> 48

<212> PRT

<213> Hepatitis E Virus

<220>

<223> M 3-2e

<400> 170

```

Thr Phe Asp Tyr Pro Gly Arg Ala His Thr Phe Asp Asp Phe Cys Pro
 1          5          10          15
Glu Cys Arg Ala Leu Gly Leu Gln Gly Cys Ala Phe Gln Ser Thr Val
          20          25          30
Ala Glu Leu Gln Arg Leu Lys Val Lys Val Gly Lys Thr Arg Glu Leu
          35          40          45

```

<210> 171

<211> 33

<212> PRT

<213> Hepatitis E Virus

<220>

<223> B 4-2

<400> 171

Ala	Asn	Pro	Pro	Asp	His	Ser	Ala	Pro	Leu	Gly	Val	Thr	Arg	Pro	Ser
1				5					10					15	
Ala	Pro	Pro	Leu	Pro	His	Val	Val	Asp	Leu	Pro	Gln	Leu	Gly	Pro	Arg
			20					25					30		

Arg

<210> 172

<211> 48

<212> PRT

<213> Hepatitis E Virus

<220>

<223> B 3-2e

<400> 172

Thr	Leu	Asp	Tyr	Pro	Ala	Arg	Ala	His	Thr	Phe	Asp	Asp	Phe	Cys	Pro
1				5					10					15	
Glu	Cys	Arg	Pro	Leu	Gly	Leu	Gln	Gly	Cys	Ala	Phe	Gln	Ser	Thr	Val
			20					25					30		
Ala	Glu	Leu	Gln	Arg	Leu	Lys	Met	Lys	Val	Gly	Lys	Thr	Arg	Glu	Leu
			35				40					45			

<210> 173

<211> 33

<212> PRT

<213> Hepatitis E Virus

<220>

<223> OFR3 (u4.2)

<400> 173

Asp	Ser	Arg	Pro	Ala	Pro	Ser	Val	Pro	Leu	Gly	Val	Thr	Ser	Pro	Ser
1				5					10					15	
Ala	Pro	Pro	Leu	Pro	Pro	Val	Val	Asp	Leu	Pro	Gln	Leu	Gly	Leu	Arg
			20					25					30		

Arg

<210> 174

<211> 48

<212> PRT

<213> Hepatitis E Virus

<220>

<223> ORF2 (u3.2e)

<400> 174

Thr	Val	Asp	Tyr	Pro	Ala	Arg	Ala	His	Thr	Phe	Asp	Asp	Phe	Cys	Pro
1				5					10					15	
Glu	Cys	Arg	Thr	Leu	Gly	Leu	Gln	Gly	Cys	Ala	Phe	Gln	Ser	Thr	Ile

[illegible]

Cys Ala Phe Gln Ser Thr Ile Ala Glu Leu Gln Arg Leu Lys Met Lys
 305 310 315 320
 Val Gly Lys Thr Arg Glu Ser
 325

<210> 176
 <211> 327
 <212> PRT
 <213> Hepatitis E Virus

<220>
 <223> US-2 SG3

<223> Xaa = Unknown or Other at position 114

<400> 176
 Gly Ala Asp Gly Thr Ala Glu Leu Thr Thr Thr Ala Ala Thr Arg Phe
 1 5 10 15
 Met Lys Asp Leu His Phe Ala Gly Thr Asn Gly Val Gly Glu Val Gly
 20 25 30
 Arg Gly Ile Ala Leu Thr Leu Phe Asn Leu Ala Asp Thr Leu Leu Gly
 35 40 45
 Gly Leu Pro Thr Glu Leu Ile Ser Ser Ala Gly Gly Gln Leu Phe Tyr
 50 55 60
 Ser Arg Pro Val Val Ser Ala Asn Gly Glu Pro Thr Val Lys Leu Tyr
 65 70 75 80
 Thr Ser Val Glu Asn Ala Gln Gln Asp Lys Gly Ile Thr Ile Pro His
 85 90 95
 Asp Ile Asp Leu Gly Asp Ser Arg Val Val Ile Gln Asp Tyr Asp Asn
 100 105 110
 Gln Xaa Glu Gln Asp Arg Pro Thr Pro Ser Pro Ala Pro Ser Arg Pro
 115 120 125
 Phe Ser Val Leu Arg Ala Asn Asp Val Leu Trp Leu Ser Leu Thr Ala
 130 135 140
 Ala Glu Tyr Asp Gln Thr Thr Tyr Gly Ser Ser Thr Asn Pro Met Tyr
 145 150 155 160
 Val Ser Asp Thr Val Thr Leu Val Asn Val Ala Thr Gly Ala Gln Ala
 165 170 175
 Val Ala Arg Ser Leu Asp Trp Ser Lys Val Thr Leu Asp Gly Arg Pro
 180 185 190
 Leu Thr Thr Ile Gln Gln Tyr Ser Lys Thr Phe Tyr Val Leu Pro Leu
 195 200 205
 Arg Gly Lys Leu Ser Phe Trp Glu Ala Gly Thr Thr Lys Ala Gly Tyr
 210 215 220
 Pro Tyr Asn Tyr Asn Thr Thr Ala Ser Asp Gln Ile Leu Ile Glu Asn
 225 230 235 240
 Ala Ala Gly His Arg Val Ala Ile Ser Thr Tyr Thr Thr Ser Leu Gly
 245 250 255
 Ala Gly Pro Thr Ser Ile Ser Ala Val Gly Val Leu Ala Pro His Ser
 260 265 270
 Ala Leu Ala Val Leu Glu Asp Thr Ile Asp Tyr Pro Ala Arg Ala His
 275 280 285
 Thr Phe Asp Asp Phe Cys Pro Glu Cys Arg Thr Leu Gly Leu Gln Gly
 290 295 300
 Cys Ala Phe Gln Ser Thr Ile Ala Glu Leu Gln Arg Leu Lys Met Lys
 305 310 315 320
 Val Gly Lys Thr Arg Glu Ser

325

<210> 177
 <211> 21
 <212> DNA
 <213> Hepatitis E Virus

<220>
 <223> HEVConsORF1-s2

<400> 177
 ctgcccytkgc gaatgctgtg g 21

<210> 178
 <211> 24
 <212> DNA
 <213> Hepatitis E Virus

<220>
 <223> HEVConsORF1-a2

<400> 178
 ggcagwrtac carcgtgaa catc 24

<210> 179
 <211> 294
 <212> DNA
 <213> Hepatitis E Virus

<220>
 <223> z12-orf1 (G.S.)

<400> 179
 tggcattact actgccattg agcaagctgc tctggctgcg gcccaattctg ccttggcgaa 60
 tgctgtgggtg gtttcggcgt ttttatctcg tttacagact gagattctta ttaatttgat 120
 gcaaccccgca cagttgggtct ttccgacctga ggtgttcttg aaccatccca tccaacgtgt 180
 tatacataat gaattggagc agtactgccg ggcccgggccc ggtcgctgtc tggaaattgg 240
 agcccatcca aggtcaatca atgataatcc taatgttctg catcggtgtt tcct 294

<210> 180
 <211> 418
 <212> DNA
 <213> Hepatitis E Virus

<220>
 <223> z12-orf1.con

<400> 180
 ctggcattac tactgctatt gagcaagctg ctctgggtgc ggccaattct gccttggcga 60
 atgctgtggt ggttcggcgt ttttatctc gtttacagac tgagattctt attaatttga 120
 tgcaaccccg acagttggtc ttccgacctg aggtgttctg gaaccatccc atccaacgtg 180
 ttatacataa tgaattggag cagtactgcc ggcccgggccc cggtcgctgt ctggaaattg 240
 gagcccatcc aaggtcaatc aatgataatc ctaatgttct gcatcggtgc tttttacgac 300
 cggtcgggag ggaagttcag cgctggtact cgcggccccc ccgtggcccc gcggccaact 360
 gccgcgggtc tgcgctgcgt ggtctccccc ctgtcgaccg cacttactgc ctcgatgg 418

<210> 181

```

<211> 197
<212> DNA
<213> Hepatitis E Virus

<220>
<223> z12-orf2.con

<400> 181
gacagaatta atttcgtcgg ctgggggtca actgtttctac tcccgcacctg togtctcagc      60
caatggcgag ccgactgtca agttatacac atctgttgag aatgcacagc aggataaggg      120
gatagctatt ccacatgaca tagatttggg cgactctcgt ttggtaatcc aggattatga      180
taaccaacac gaacaag                                     197

<210> 182
<211> 25
<212> DNA
<213> Hepatitis E Virus

<220>
<223> HEVConsORF2/3-s1

<400> 182
gtatcgggkyk gaatgaataa catgt                                     25

<210> 183
<211> 25
<212> DNA
<213> Hepatitis E Virus

<220>
<223> HEVConsORF2/3-a1

<400> 183
agggggttggg tggatgaata taggg                                     25

<210> 184
<211> 234
<212> DNA
<213> Hepatitis E Virus

<220>
<223> z12.orf23.con

<400> 184
gtatcgggktt gaatgaataa catgttttgt gcatcgccca tgggatcacc atgcgcacct      60
gggttggttct gttgttggtc ctctgtgttc tgccatagct gcccgcgcca ccggccggcc      120
agycgactgg ccgccgtcgt gggcgggcga gggcggtgc cggcggtggt ttctgggggtg      180
acagggttga ttctcagccc ttccgacctc cctatatattca tccaaccaac ccct      234

<210> 185
<211> 890
<212> DNA
<213> Hepatitis E Virus

<220>
<223> z12-3p.race

```

```

<400> 185
gtcgtctcgg ccaatggcga gccgactgtc aagttataca catctgttga gaatgcacag      60
caggataaagg ggatagctat tccacatgac atagatttgg gcgactctcg ttgggtaatc      120
caggattacg ataatcagca cgagcaggac cggcccaccc ctccgcccgc cccgtctcgt      180
cctttctcgg tctccgcgc taatgatgct ttgtggcttt ctcttacgcg tgcctgagtat      240
gaccagacta catatgggtc gtccaccaac ccgatgtatg tctcagacac tgttacattt      300
gtcaatgtgg ccacaggggc tcaggctgtc gcccgttctc ttgattggto taaagttacc      360
ctggacgggc gccctcttac taccatccag cagtactcta agacatttta tgttctccca      420
cttcgoggga agttatcttt ttgggaggct ggcacaacta aagccggtta cccttataat      480
tataacacaa ctgctagtga ccagattctg attgaaaacg cggctggcca tctgtctcgt      540
atatctactt atactactag cctgggcgcc ggccctgtgt cagtttctgc ggttggtgtg      600
ttagcccccac actcagacct tgctattctt gaagacactg ttgactatcc ggcccggtgt      660
cacacttttg atgacttctg tccggaatgc cgtgccctgg gtctgcaggg gtgtgctttt      720
caatctacta tgcctgagct ccagcgtctt aaaatgaagg taggcaaac ccgggagttt      780
taattaattc ttcttgtgcc ccttcacagg ttctcgcttt atttctttct tctgcctccc      840
gcgctccctg gaaaaaaaaa aaaaaaaaaa gtactagtgc acgctgggcc      890

```

```

<210> 186
<211> 919
<212> DNA
<213> Hepatitis E Virus

```

```

<220>
<223> z12-3p.con

```

```

<400> 186
gacagaatta atttcgtcgg ctgggggtca actgtttctac tcccgcctcg tcgtctcagc      60
caatggcgag ccgactgtca agttatacac atctgttgag aatgcacagc aggataaggg      120
gatagctatt ccacatgaca tagatttggg cgactctcgt ttggtaatcc aggattacga      180
taatcagcac gagcaggacc ggcccacccc ttccgcccgc cgtctcgtc ctttctcggg      240
cctccgcgct aatgatgctt tgtggctttc tcttaccgct gctgagtatg accagactac      300
atatgggtcg tccaccaacc cgaatgtatg tctcagacact gttacatttg tcaatgtggc      360
cacaggggct caggctgtcg cccgttctct tgattgggtc aaagttaccc tggacggccg      420
ccctcttact accatccagc agtaactctaa gacattttat gttctccac ttccggggaa      480
gttatctttt tgggaggctg gcacaactaa agccggttac cttataaatt ataacacaa      540
tgctagtga cagattctga ttgaaacgc ggctggccat cgtgtcgcta tatctactta      600
tactactagc ctgggcgcgc gccctgtgtc agtttctgcg gttggtgtgt tagcccaca      660
ctcgagcctt gctattcttg aagacactgt tgactatccg gcccggtgtc acacttttga      720
tgacttctgt ccggaatgcc gtgccctggg tctgcagggg tgtgcttttc aatctactat      780
cgctgagctc cagcgtctta aaatgaaggg aggcacaaac cgggagtttt aatataattc      840
tcttgtgccc ccttcacggg tctcgcttta tttctttctt ctgcctcccg cgtccctggg      900
aaaaaaaaa aaaaaaaaaa

```

```

<210> 187
<211> 138
<212> PRT
<213> Hepatitis E Virus

```

```

<220>
<223> z12-orf1.pep

```

```

<400> 187
Gly Ile Thr Thr Ala Ile Glu Gln Ala Ala Leu Gly Ala Ala Asn Ser
 1             5             10             15
Ala Leu Ala Asn Ala Val Val Val Arg Pro Phe Leu Ser Arg Leu Gln
                20             25             30
Thr Glu Ile Leu Ile Asn Leu Met Gln Pro Arg Gln Leu Val Phe Arg

```

```

      35          40          45
Pro Glu Val Phe Trp Asn His Pro Ile Gln Arg Val Ile His Asn Glu
  50          55          60
Leu Glu Gln Tyr Cys Arg Ala Arg Ala Gly Arg Cys Leu Glu Ile Gly
  65          70          75          80
Ala His Pro Arg Ser Ile Asn Asp Asn Pro Asn Val Leu His Arg Cys
      85          90          95
Phe Leu Arg Pro Val Gly Arg Asp Val Gln Arg Trp Tyr Ser Ala Pro
      100          105          110
Thr Arg Gly Pro Ala Ala Asn Cys Arg Arg Ser Ala Leu Arg Gly Leu
      115          120          125
Pro Pro Val Asp Arg Thr Tyr Cys Leu Asp
      130          135

```

<210> 188
 <211> 61
 <212> PRT
 <213> Hepatitis E Virus

<220>
 <223> z12-orf2-5'.pep

<223> Xaa = Unknown or Other at position 25

```

      <400> 188
Met Arg Pro Arg Val Val Leu Leu Leu Phe Leu Val Phe Leu Pro Met
  1          5          10          15
Leu Pro Ala Pro Pro Ala Gly Gln Xaa Thr Gly Arg Arg Arg Gly Arg
      20          25          30
Arg Ser Gly Gly Ala Gly Gly Gly Phe Trp Gly Asp Arg Val Asp Ser
      35          40          45
Gln Pro Phe Ala Leu Pro Tyr Ile His Pro Thr Asn Pro
      50          55          60

```

<210> 189
 <211> 276
 <212> PRT
 <213> Hepatitis E Virus

<220>
 <223> z12-orf2-3'.pep

```

      <400> 189
Thr Glu Leu Ile Ser Ser Ala Gly Gly Gln Leu Phe Tyr Ser Arg Pro
  1          5          10          15
Val Val Ser Ala Asn Gly Glu Pro Thr Val Lys Leu Tyr Thr Ser Val
      20          25          30
Glu Asn Ala Gln Gln Asp Lys Gly Ile Ala Ile Pro His Asp Ile Asp
      35          40          45
Leu Gly Asp Ser Arg Leu Val Ile Gln Asp Tyr Asp Asn Gln His Glu
      50          55          60
Gln Asp Arg Pro Thr Pro Ser Pro Ala Pro Ser Arg Pro Phe Ser Val
      65          70          75          80
Leu Arg Ala Asn Asp Ala Leu Trp Leu Ser Leu Thr Ala Ala Glu Tyr
      85          90          95
Asp Gln Thr Thr Tyr Gly Ser Ser Thr Asn Pro Met Tyr Val Ser Asp
      100          105          110

```

```

Thr Val Thr Phe Val Asn Val Ala Thr Gly Ala Gln Ala Val Ala Arg
      115              120              125
Ser Leu Asp Trp Ser Lys Val Thr Leu Asp Gly Arg Pro Leu Thr Thr
      130              135              140
Ile Gln Gln Tyr Ser Lys Thr Phe Tyr Val Leu Pro Leu Arg Gly Lys
145              150              155              160
Leu Ser Phe Trp Glu Ala Gly Thr Thr Lys Ala Gly Tyr Pro Tyr Asn
      165              170              175
Tyr Asn Thr Thr Ala Ser Asp Gln Ile Leu Ile Glu Asn Ala Ala Gly
      180              185              190
His Arg Val Ala Ile Ser Thr Tyr Thr Thr Ser Leu Gly Ala Gly Pro
      195              200              205
Val Ser Val Ser Ala Val Gly Val Leu Ala Pro His Ser Ser Leu Ala
      210              215              220
Ile Leu Glu Asp Thr Val Asp Tyr Pro Ala Arg Ala His Thr Phe Asp
225              230              235              240
Asp Phe Cys Pro Glu Cys Arg Ala Leu Gly Leu Gln Gly Cys Ala Phe
      245              250              255
Gln Ser Thr Ile Ala Glu Leu Gln Arg Leu Lys Met Lys Val Gly Lys
      260              265              270
Thr Arg Glu Phe
      275

```

<210> 190
 <211> 74
 <212> PRT
 <213> Hepatitis E Virus

<220>
 <223> z12-orf3.pep

```

      <400> 190
Met Asn Asn Met Phe Cys Ala Ser Pro Met Gly Ser Pro Cys Ala Leu
 1      5      10      15
Gly Leu Phe Cys Cys Cys Ser Ser Cys Phe Cys Leu Cys Cys Pro Arg
      20      25      30
His Arg Pro Ala Ser Arg Leu Ala Val Val Gly Gly Ala Ala Ala
      35      40      45
Val Pro Ala Val Val Ser Gly Val Thr Gly Leu Ile Leu Ser Pro Ser
      50      55      60
Pro Ser Pro Ile Phe Ile Gln Pro Thr Pro
65              70

```

<210> 191
 <211> 408
 <212> DNA
 <213> Hepatitis E Virus

<220>
 <223> pJOorf3-29.seq

```

      <400> 191
gaattcatga ataacatgtc ttttgcacgc cccatgggat caccatgcgc cctagggctg      60
ttctgttggt gctcttcgtg cttttgccta tgetgccgcg gccaccggcc agccagccgt      120
ctggccgcgc tcgtgggcgc cgcagcggcg gtgccggcgc tggtttctgg ggtgacaggg      180
ttgattctca gcccttcgcc ctccctata ttcattcaac caaccccttc gccgccgatg      240
tcgtttcaca acccggggct ggaactcgcc ctgcacagcc gcccgcgcc cttggctccg      300

```

```

cttggcgtga ccagtcaccag cgcacctccg ctgccccccg tcgtcgatct gccccagctt 360
ggtctgcgcc gcgactacaa ggacgacgat gacaagtaat aaggatcc 408

```

```

<210> 192
<211> 1026
<212> DNA
<213> Hepatitis E Virus

```

```

<220>
<223> cksorf2m-2.seq

```

```

<400> 192
gaattcatgg gtgctgatgg gactgctgag cttactacca cagcagccac acgtttcatg 60
aaggacctgc acttcgctgg cacgaatggc gttggtgagg tgggtcgtgg tatcgccctg 120
acactgtttca atctcgctga tacgcttctc ggcggtttac cgacagaatt gatttcgtcg 180
gctggggggcc aactgtttta ctcccggccg gttgtctcag ccaatggcga gccaaacagta 240
aagtatatata catctgttga gaatgcgcag caagacaagg gcatcaccat tccacatgat 300
atagacctgg gtgactcccg tgtggttacc caggattatg ataaccagca tgagcaagac 360
cgacctactc cgtcacctgc cccctctcgc ccttctcag ttcttcgtgc caatgatgtt 420
ttgtggcttt cctcactgc cgtgagtat gaccagacta cgtatgggtc gtccaccaac 480
cctatgtatg tctctgacac agttacgctt gttaatgtgg ctactggtgc tcaggctgtt 540
gcccgctccc ttgattggtc taaagttaact ctggacggcc gcccccttac taccattcag 600
cagtattcta agacatttta tgttctcccg ctccgcgga agctgtcctt ttgggaggct 660
ggcacgacta aggcgggcta ccttacaat tataatacta ccgctagtga ccaaattttg 720
attgagaatg cggccggcca cgtgtcgtc atttccacct ataccactag cttaggtgcc 780
ggtcctacct cgatctctgc ggtcggcgta ctggctccac actctgccct tgccgttctt 840
gaggatacta ttgattaccc cgcctgtgcc catacttttg atgatttttg ccgggagtgc 900
cgtaccctag gtttcagggt ttgtgcattc cagtctacta ttgctgagct ccagcgttta 960
aaaatgaagg taggtaaaac cggggagtct gactacaagg acgacgatga caagtaataa 1020
ggatcc 1026

```

```

<210> 193
<211> 1389
<212> DNA
<213> Hepatitis E Virus

```

```

<220>
<223> CKSORF32M-3.seq

```

```

<400> 193
gaattcatga ataactgtc ttttgcacgc cccatgggat caccatgcgc cctagggctg 60
ttctgttggt gctcttcgtg cttttgccta tgcgtcccg cccacgggac agccagccgt 120
ctggccggcg tcgtggggcg cgtagcggcg gtgcccggcg tggtttctgg ggtgacaggg 180
ttgattctca gcccttcgcc ctcccctata ttcatccaac caaccccttc gccgccgatg 240
tcgtttcaca acccggggct ggaactcgcc ctgcacagcc gcccgcgcc cttgggtccg 300
cttggcgtga ccagtcaccg cgcacctccg ctgccccccg tcgtcgatct gccccagctt 360
ggtctgcgcc gcggtgctga tgggactgct gagcttacta ccacagcagc cacacgtttc 420
atgaaggacc tgcacttcgc tggcacgaat ggcgttggtg aggtgggtcg tggtatcgcc 480
ctgacactgt tcaatctcgc tgatacgctt ctggcggtt taccgacaga attgatttcg 540
tcggctgggg gccaaactgt ttaactccgc ccggttgtct cagccaatgg cgagccaaca 600
gtaaagttat atacatctgt tgagaatgcg cagcaagaca agggcatcac cattccacat 660
gatataagacc tgggtgactc ccgtgtggtt atccaggatt atgataacca gcatgagcaa 720
gaccgaccta ctccgtcacc tgcccctctc cgccttctc cagttcttcg tgccaatgat 780
gttttgtggc ttccctcacc tgccgctgag tatgaccaga ctacgtatgg gtcgtccacc 840
aaccctatgt atgtctctga cacagttacg cttgttaatg tggctactgg tgctcaggct 900
gttgcgcgct cccttgattg gtctaaagt actctggacg gccgccccct tactaccatt 960
cagcagttat ctaagacatt ttatgttctc ccgctccgcg ggaagctgtc cttttgggag 1020

```

```

gctggcacga ctaaggccgg ctacccttac aattataata ctaccgctag tgaccaaatt 1080
ttgattgaga atgogggccgg ccacogtggtc gctatttoca cctataccac tagcttaggt 1140
gccggtccta cctcgatctc tgcggtcggc gtactggctc cacactctgc ccttgccgtt 1200
cttgaggata ctattgatta ccccgcccggt gcccatactt ttgatgattt ttgcccggag 1260
tgcogtaacc taggttttga gggttgtgca ttccagtcta ctattgctga gctccagcgt 1320
ttaaaaatga aggtaggtaa aacccgggag tctgactaca aggacgacga tgacaagtaa 1380
taaggatcc 1389

```

```

<210> 194
<211> 408
<212> DNA
<213> Hepatitis E Virus

```

```

<220>
<223> plorf3-12.con

```

```

<400> 194
gaattcatga ataacatgtc ttttgcacgc cccatgggat caccatgcgc cctagggctg 60
ttctgttggt gctcttcgtg cttttgccta tgctgccgc gccacggcc gccacggct 120
ctggccggcg tgcgtggggc cgcagcggcg gtgcggggc tggtttctgg ggtgacaggg 180
ttgattctca gcccttcgca ctccctata ttcacccaac caaccccttc gccgccgatg 240
tcgtttcaca acccggggct ggaactcgcc ctgcacagcc gcccgcggcc cttggctccg 300
cttggcgatga ccagtcacag cgcctctcgc ctgcccccg tcgtcgatct gccccagctt 360
ggtctgcgcc gcgactacaa ggacgacgat gacaagtaat aaggatcc 408

```

```

<210> 195
<211> 1026
<212> DNA
<213> Hepatitis E Virus

```

```

<220>
<223> plorf2.2-6.seq

```

```

<400> 195
gaattcatgg gtgctgatgg gactgctgag cttactacca cagcagccac acgtttcatg 60
aaggacctgc acttcgctgg cacgaatggc gttggtgagg tgggtcgtgg tatcgccctg 120
acactgttca atctogctga taogcttctc ggoggtttac cgacagaatt gatttcgtcg 180
gctggggggc aactgtttta ctccgcggcg gttgtctcag ccaatggcga gccaacagta 240
aagttatata catctgttga gaatgcgcag caagacaagg gcatcaccat tccacatgat 300
atagacctgg gtgactcccg tgtggttatc caggattatg ataaccagca tgagcaagac 360
cgacctactc cgtcacctgc cccctctcgc ccttctcag ttcttcgtgc caatgatgtt 420
ttgtggcttt cctcactgc cgtgagtat gaccagacta cgtatgggtc gtccaccaac 480
cctatgtatg tctctgacac agttacgctt gttaatgtgg ctactgggtc tcaggctgtt 540
gcccgctccc ttgattggtc taaagttact ctggacggcc gccccttac taccattcag 600
cagtattcta agacatttta tgttctcccg ctccgcggga agctgtcctt ttgggaggct 660
ggcacgacta aggcgggcta cccttacaat tataatacta ccgctagtga ccaaattttg 720
attgagaatg cggccggcca ccgtgtcgct atttccacct ataccactag cttaggtgac 780
ggtcctacct cgatctctgc ggtcggcgta ctggctccac actctgccct tgcggttctt 840
gaggatacta ttgattaccc cgcctgtgcc catacttttg atgatttttg cccggagtcg 900
cgtaccctag gtttgcaggg ttgtgcattc cagtctacta ttgctgagct ccagcgttta 960
aaaatgaagg taggtaaaac ccgggagctc gactacaagg acgacgatga caagtaataa 1020
ggatcc 1026

```

```

<210> 196
<211> 1389
<212> DNA
<213> Hepatitis E Virus

```

<220>

<223> PLORF32M-14-5.seq

<400> 196

```

gaattcatga ataacatgtc ttttgcacgc cccatgggat caccatgcgc cctagggctg      60
ttctgttggt gctcttcgtg cttttgccta tgctgcccgc gccacgggcc agccagccgt      120
ctggccgccc tctgtggcgg cgtagcggcg gtgcccggcg tggtttcttg ggtgacaggg      180
ttgattctca gcccttcgcc ctccctata ttcatccaac caaccccttc gccgcccgat      240
tcgtttcaca acccggggct ggaactcgcc ctgcacagcc gcccggcccc cttgggtccc      300
cttggcgctga ccagtcccag cgcctcccg ctgcccccg tcgtcgatct gccccagctt      360
ggtctgcgcc gcggtgctga tgggactgct gagcttacta ccacagcagc cacacgtttc      420
atgaaggacc tgcacttcgc tggcacgaat ggcgttggtg aggtgggtcg tggtatcgcc      480
ctgacactgt tcaatctcgc tgatacgctt ctggcggtt taccgacaga attgatttcg      540
tcggctgggg gccaaactgtt ttactcccg cgggttgtct cagccaatgg cgagccaaca      600
gtaaagttat atacatctgt tgagaatgcg cagcaagaca agggcatcac cattccacat      660
gatatagacc tgggtgactc ccgtgtggtt atccaggatt atgataacca gcatgagcaa      720
gaccgaccta ctccgtcacc tgccccctct cgcctctct cagttcttcg tgccaatgat      780
gttttgtggc tttccctcac tgccgctgag tatgaccaga ctacgtatgg gtcgtccacc      840
aaccctatgt atgtctctga cacagttacg cttgttaatg tggctactgg tgctcaggct      900
gttgcctcgt cctctgattg gtctaaagt actctggacg gccgccccct tactaccatt      960
cagcagttat ctaagacatt ttatgttctc ccgctccgcg ggaagctgtc cttttgggag     1020
gctggcacga ctaaggcccg ctacccttac aattataata ctaccgctag tgaccaaatt     1080
ttgattgaga atgcggcccg ccaccgtgtc gctatttcca cctataccac tagcttaggt     1140
gccggtccta cctcgatctc tgccgtcgcc gtactggctc cacactctgc cttgcccgtt     1200
cttgaggata ctattgatta cccgcgccgt gccatactt ttgatgattt ttgcccggag     1260
tgccgtaccc taggtttgca gggttgtgca ttccagtcta ctattgctga gctccagcgt     1320
ttaaaaatga aggtaggtaa aaccggggag tctgactaca aggacgacga tgacaagtaa     1380
taagatccc

```

<210> 197

<211> 74

<212> PRT

<213> Hepatitis E Virus

<220>

<223> z12-orf3-5'.pep

<223> Xaa = Unknown or Other at position 37

<400> 197

```

Met Asn Asn Met Phe Cys Ala Ser Pro Met Gly Ser Pro Cys Ala Leu
 1          5          10          15
Gly Leu Phe Cys Cys Cys Ser Ser Cys Phe Cys Leu Cys Cys Pro Arg
          20          25          30
His Arg Pro Ala Xaa Arg Leu Ala Val Val Gly Gly Ala Ala Ala
          35          40          45
Val Pro Ala Val Val Ser Gly Val Thr Gly Leu Ile Leu Ser Pro Ser
          50          55          60
Pro Ser Pro Ile Phe Ile Gln Pro Thr Pro
65          70

```

<210> 198

<211> 63

<212> DNA

<213> Hepatitis E Virus

<220>

<223> Primer orf23p

<400> 198

tatatggatc cttattactt gtcacgtctg tccttgtagt cagactcccg ggttttacct
acc

60

63

<210> 199

<211> 338

<212> PRT

<213> Hepatitis E Virus

<220>

<223> cksorf2m-2.pep

<400> 199

Glu Phe Met Gly Ala Asp Gly Thr Ala Glu Leu Thr Thr Thr Ala Ala
1 5 10 15
Thr Arg Phe Met Lys Asp Leu His Phe Ala Gly Thr Asn Gly Val Gly
20 25 30
Glu Val Gly Arg Gly Ile Ala Leu Thr Leu Phe Asn Leu Ala Asp Thr
35 40 45
Leu Leu Gly Gly Leu Pro Thr Glu Leu Ile Ser Ser Ala Gly Gly Gln
50 55 60
Leu Phe Tyr Ser Arg Pro Val Val Ser Ala Asn Gly Glu Pro Thr Val
65 70 75 80
Lys Leu Tyr Thr Ser Val Glu Asn Ala Gln Gln Asp Lys Gly Ile Thr
85 90 95
Ile Pro His Asp Ile Asp Leu Gly Asp Ser Arg Val Val Ile Gln Asp
100 105 110
Tyr Asp Asn Gln His Glu Gln Asp Arg Pro Thr Pro Ser Pro Ala Pro
115 120 125
Ser Arg Pro Phe Ser Val Leu Arg Ala Asn Asp Val Leu Trp Leu Ser
130 135 140
Leu Thr Ala Ala Glu Tyr Asp Gln Thr Thr Tyr Gly Ser Ser Thr Asn
145 150 155 160
Pro Met Tyr Val Ser Asp Thr Val Thr Leu Val Asn Val Ala Thr Gly
165 170 175
Ala Gln Ala Val Ala Arg Ser Leu Asp Trp Ser Lys Val Thr Leu Asp
180 185 190
Gly Arg Pro Leu Thr Thr Ile Gln Gln Tyr Ser Lys Thr Phe Tyr Val
195 200 205
Leu Pro Leu Arg Gly Lys Leu Ser Phe Trp Glu Ala Gly Thr Thr Lys
210 215 220
Ala Gly Tyr Pro Tyr Asn Tyr Asn Thr Thr Ala Ser Asp Gln Ile Leu
225 230 235 240
Ile Glu Asn Ala Ala Gly His Arg Val Ala Ile Ser Thr Tyr Thr Thr
245 250 255
Ser Leu Gly Ala Gly Pro Thr Ser Ile Ser Ala Val Gly Val Leu Ala
260 265 270
Pro His Ser Ala Leu Ala Val Leu Glu Asp Thr Ile Asp Tyr Pro Ala
275 280 285
Arg Ala His Thr Phe Asp Asp Phe Cys Pro Glu Cys Arg Thr Leu Gly
290 295 300
Leu Gln Gly Cys Ala Phe Gln Ser Thr Ile Ala Glu Leu Gln Arg Leu
305 310 315 320
Lys Met Lys Val Gly Lys Thr Arg Glu Ser Asp Tyr Lys Asp Asp Asp

				325					330					335
Asp Lys														
<210>	200													
<211>	338													
<212>	PRT													
<213>	Hepatitis E Virus													
<220>														
<223>	plorf2.2-6.pep													
<400>	200													
Glu Phe Met Gly Ala Asp Gly Thr Ala Glu Leu Thr Thr Thr Ala Ala														
1 5 10 15														
Thr Arg Phe Met Lys Asp Leu His Phe Ala Gly Thr Asn Gly Val Gly														
20 25 30														
Glu Val Gly Arg Gly Ile Ala Leu Thr Leu Phe Asn Leu Ala Asp Thr														
35 40 45														
Leu Leu Gly Gly Leu Pro Thr Glu Leu Ile Ser Ser Ala Gly Gly Gln														
50 55 60														
Leu Phe Tyr Ser Arg Pro Val Val Ser Ala Asn Gly Glu Pro Thr Val														
65 70 75 80														
Lys Leu Tyr Thr Ser Val Glu Asn Ala Gln Gln Asp Lys Gly Ile Thr														
85 90 95														
Ile Pro His Asp Ile Asp Leu Gly Asp Ser Arg Val Val Ile Gln Asp														
100 105 110														
Tyr Asp Asn Gln His Glu Gln Asp Arg Pro Thr Pro Ser Pro Ala Pro														
115 120 125														
Ser Arg Pro Phe Ser Val Leu Arg Ala Asn Asp Val Leu Trp Leu Ser														
130 135 140														
Leu Thr Ala Ala Glu Tyr Asp Gln Thr Thr Tyr Gly Ser Ser Thr Asn														
145 150 155 160														
Pro Met Tyr Val Ser Asp Thr Val Thr Leu Val Asn Val Ala Thr Gly														
165 170 175														
Ala Gln Ala Val Ala Arg Ser Leu Asp Trp Ser Lys Val Thr Leu Asp														
180 185 190														
Gly Arg Pro Leu Thr Thr Ile Gln Gln Tyr Ser Lys Thr Phe Tyr Val														
195 200 205														
Leu Pro Leu Arg Gly Lys Leu Ser Phe Trp Glu Ala Gly Thr Thr Lys														
210 215 220														
Ala Gly Tyr Pro Tyr Asn Tyr Asn Thr Thr Ala Ser Asp Gln Ile Leu														
225 230 235 240														
Ile Glu Asn Ala Ala Gly His Arg Val Ala Ile Ser Thr Tyr Thr Thr														
245 250 255														
Ser Leu Gly Ala Gly Pro Thr Ser Ile Ser Ala Val Gly Val Leu Ala														
260 265 270														
Pro His Ser Ala Leu Ala Val Leu Glu Asp Thr Ile Asp Tyr Pro Ala														
275 280 285														
Arg Ala His Thr Phe Asp Asp Phe Cys Pro Glu Cys Arg Thr Leu Gly														
290 295 300														
Leu Gln Gly Cys Ala Phe Gln Ser Thr Ile Ala Glu Leu Gln Arg Leu														
305 310 315 320														
Lys Met Lys Val Gly Lys Thr Arg Glu Ser Asp Tyr Lys Asp Asp Asp														
325 330 335														
Asp Lys														

```

<210> 201
<211> 37
<212> DNA
<213> Hepatitis E Virus

<220>
<223> Primer orf35p

<400> 201
tatatgaatt catgaataac atgtcttttg catcgcc          37

<210> 202
<211> 68
<212> DNA
<213> Hepatitis E Virus

<220>
<223> Primer orf33p

<400> 202
tatatggatc cttattactt gtcacgtcgc tccttgtagt cgcggcgag accaagctgg      60
ggcagatc                                     68

<210> 203
<211> 132
<212> PRT
<213> Hepatitis E Virus

<220>
<223> pJOorf3-29.pep

<400> 203
Glu Phe Met Asn Asn Met Ser Phe Ala Ser Pro Met Gly Ser Pro Cys
 1           5           10          15
Ala Leu Gly Leu Phe Cys Cys Cys Ser Ser Cys Phe Cys Leu Cys Cys
 20          25          30
Pro Arg His Arg Pro Ala Ser Arg Leu Ala Ala Val Val Gly Gly Ala
 35          40          45
Ala Ala Val Pro Ala Val Val Ser Gly Val Thr Gly Leu Ile Leu Ser
 50          55          60
Pro Ser Pro Ser Pro Ile Phe Ile Gln Pro Thr Pro Ser Pro Pro Met
 65          70          75          80
Ser Phe His Asn Pro Gly Leu Glu Leu Ala Leu Asp Ser Arg Pro Ala
 85          90          95
Pro Leu Ala Pro Leu Gly Val Thr Ser Pro Ser Ala Pro Pro Leu Pro
100         105         110
Pro Val Val Asp Leu Pro Gln Leu Gly Leu Arg Arg Asp Tyr Lys Asp
115         120         125
Asp Asp Asp Lys
130

<210> 204
<211> 132
<212> PRT
<213> Hepatitis E Virus

```

<220>

<223> plorf3-12.pep

<400> 204

Glu Phe Met Asn Asn Met Ser Phe Ala Ser Pro Met Gly Ser Pro Cys
 1 5 10 15
 Ala Leu Gly Leu Phe Cys Cys Cys Ser Ser Cys Phe Cys Leu Cys Cys
 20 25 30
 Pro Arg His Arg Pro Ala Ser Arg Leu Ala Ala Val Val Gly Gly Ala
 35 40 45
 Ala Ala Val Pro Ala Val Val Ser Gly Val Thr Gly Leu Ile Leu Ser
 50 55 60
 Pro Ser Pro Ser Pro Ile Phe Ile Gln Pro Thr Pro Ser Pro Pro Met
 65 70 75 80
 Ser Phe His Asn Pro Gly Leu Glu Leu Ala Leu Asp Ser Arg Pro Ala
 85 90 95
 Pro Leu Ala Pro Leu Gly Val Thr Ser Pro Ser Ala Pro Pro Leu Pro
 100 105 110
 Pro Val Val Asp Leu Pro Gln Leu Gly Leu Arg Arg Asp Tyr Lys Asp
 115 120 125
 Asp Asp Asp Lys
 130

<210> 205

<211> 48

<212> DNA

<213> Hepatitis E Virus

<220>

<223> Primer orf23

<400> 205

ctcagcagtc ccatcagcac cgcggcgcag accaagctgg ggcagatc

48

<210> 206

<211> 459

<212> PRT

<213> Hepatitis E Virus

<220>

<223> CKSORF32M-3.pep

<400> 206

Glu Phe Met Asn Asn Met Ser Phe Ala Ser Pro Met Gly Ser Pro Cys
 1 5 10 15
 Ala Leu Gly Leu Phe Cys Cys Cys Ser Ser Cys Phe Cys Leu Cys Cys
 20 25 30
 Pro Arg His Arg Pro Ala Ser Arg Leu Ala Ala Val Val Gly Gly Val
 35 40 45
 Ala Ala Val Pro Ala Val Val Ser Gly Val Thr Gly Leu Ile Leu Ser
 50 55 60
 Pro Ser Pro Ser Pro Ile Phe Ile Gln Pro Thr Pro Ser Pro Pro Met
 65 70 75 80
 Ser Phe His Asn Pro Gly Leu Glu Leu Ala Leu Asp Ser Arg Pro Ala
 85 90 95
 Pro Leu Ala Pro Leu Gly Val Thr Ser Pro Ser Ala Pro Pro Leu Pro
 100 105 110

```

Pro Val Val Asp Leu Pro Gln Leu Gly Leu Arg Arg Gly Ala Asp Gly
      115      120      125
Thr Ala Glu Leu Thr Thr Thr Ala Ala Thr Arg Phe Met Lys Asp Leu
      130      135      140
His Phe Ala Gly Thr Asn Gly Val Gly Glu Val Gly Arg Gly Ile Ala
145      150      155      160
Leu Thr Leu Phe Asn Leu Ala Asp Thr Leu Leu Gly Gly Leu Pro Thr
      165      170      175
Glu Leu Ile Ser Ser Ala Gly Gly Gln Leu Phe Tyr Ser Arg Pro Val
      180      185      190
Val Ser Ala Asn Gly Glu Pro Thr Val Lys Leu Tyr Thr Ser Val Glu
      195      200      205
Asn Ala Gln Gln Asp Lys Gly Ile Thr Ile Pro His Asp Ile Asp Leu
      210      215      220
Gly Asp Ser Arg Val Val Ile Gln Asp Tyr Asp Asn Gln His Glu Gln
225      230      235      240
Asp Arg Pro Thr Pro Ser Pro Ala Pro Ser Arg Pro Phe Ser Val Leu
      245      250      255
Arg Ala Asn Asp Val Leu Trp Leu Ser Leu Thr Ala Ala Glu Tyr Asp
      260      265      270
Gln Thr Thr Tyr Gly Ser Ser Thr Asn Pro Met Tyr Val Ser Asp Thr
      275      280      285
Val Thr Leu Val Asn Val Ala Thr Gly Ala Gln Ala Val Ala Arg Ser
      290      295      300
Leu Asp Trp Ser Lys Val Thr Leu Asp Gly Arg Pro Leu Thr Thr Ile
305      310      315      320
Gln Gln Tyr Ser Lys Thr Phe Tyr Val Leu Pro Leu Arg Gly Lys Leu
      325      330      335
Ser Phe Trp Glu Ala Gly Thr Thr Lys Ala Gly Tyr Pro Tyr Asn Tyr
      340      345      350
Asn Thr Thr Ala Ser Asp Gln Ile Leu Ile Glu Asn Ala Ala Gly His
      355      360      365
Arg Val Ala Ile Ser Thr Tyr Thr Thr Ser Leu Gly Ala Gly Pro Thr
      370      375      380
Ser Ile Ser Ala Val Gly Val Leu Ala Pro His Ser Ala Leu Ala Val
385      390      395      400
Leu Glu Asp Thr Ile Asp Tyr Pro Ala Arg Ala His Thr Phe Asp Asp
      405      410      415
Phe Cys Pro Glu Cys Arg Thr Leu Gly Leu Gln Gly Cys Ala Phe Gln
      420      425      430
Ser Thr Ile Ala Glu Leu Gln Arg Leu Lys Met Lys Val Gly Lys Thr
      435      440      445
Arg Glu Ser Asp Tyr Lys Asp Asp Asp Asp Lys
450      455

```

```

<210> 207
<211> 459
<212> PRT
<213> Hepatitis E Virus

```

```

<220>
<223> FLORF32M-14-5.pep

```

```

<400> 207
Glu Phe Met Asn Asn Met Ser Phe Ala Ser Pro Met Gly Ser Pro Cys
 1      5      10      15
Ala Leu Gly Leu Phe Cys Cys Cys Ser Ser Cys Phe Cys Leu Cys Cys

```

<210> 208

<211> 36
 <212> DNA
 <213> Hepatitis E Virus

 <220>
 <223> Primer orf2mid5p

 <400> 208
 tatatgaatt catgggtgct gatgggactg ctgagc 36

 <210> 209
 <211> 418
 <212> DNA
 <213> Hepatitis E Virus

 <220>
 <223> 1440o1.seq

 <221> CDS
 <222> (3)...(416)

 <223> Xaa = Unknown or Other at position 2

 <223> Xaa = Unknown or Other at position 5

 <223> Xaa = Unknown or Other at position 137

 <400> 209
 ct ggc aty act act gcy att gag cag gct gct ctg gct gcg gcc aat 47
 Gly Xaa Thr Thr Xaa Ile Glu Gln Ala Ala Leu Ala Ala Ala Asn
 1 5 10 15

 tcc gcc ttg gcg aat gct gtg gtg gtt cgg ccg ttt tta tcc cgt gtt 95
 Ser Ala Leu Ala Asn Ala Val Val Val Arg Pro Phe Leu Ser Arg Val
 20 25 30

 caa act gat atc ctt att aac ctg atg caa ccc cgt cag ctt gtg ttc 143
 Gln Thr Asp Ile Leu Ile Asn Leu Met Gln Pro Arg Gln Leu Val Phe
 35 40 45

 cgg cct gaa gtt ctc tgg aac cat ccg atc cag cga gtt ata cat aat 191
 Arg Pro Glu Val Leu Trp Asn His Pro Ile Gln Arg Val Ile His Asn
 50 55 60

 gag ctg gaa caa tac tgt cga gcc cgc gct ggc cgc tgt ctt gag gtg 239
 Glu Leu Glu Gln Tyr Cys Arg Ala Arg Ala Gly Arg Cys Leu Glu Val
 65 70 75

 ggc gct cac cca agg tct att aat gat aac ccc aat gtt ctg cac cgg 287
 Gly Ala His Pro Arg Ser Ile Asn Asp Asn Pro Asn Val Leu His Arg
 80 85 90 95

 tgc ttt ctc cgc ccg gtt ggg aga gac gtc cag cgc tgg tat tcc gcc 335
 Cys Phe Leu Arg Pro Val Gly Arg Asp Val Gln Arg Trp Tyr Ser Ala
 100 105 110

 ccc act cgt ggt cca gcg gct aac tgc cgc cgt tct gcg cta cgc ggt 383

Pro Thr Arg Gly Pro Ala Ala Asn Cys Arg Arg Ser Ala Leu Arg Gly
 115 120 125

ttg ccc cct gtc gac cgc act tac tgt yty gat gg
 Leu Pro Pro Val Asp Arg Thr Tyr Cys Xaa Asp
 130 135

418

<210> 210
 <211> 138
 <212> PRT
 <213> Hepatitis E Virus

<220>
 <223> Xaa = Unknown or Other at position 2
 <223> Xaa = Unknown or Other at position 5
 <223> Xaa = Unknown or Other at position 137

<400> 210
 Gly Xaa Thr Thr Xaa Ile Glu Gln Ala Ala Leu Ala Ala Ala Asn Ser
 1 5 10 15
 Ala Leu Ala Asn Ala Val Val Val Arg Pro Phe Leu Ser Arg Val Gln
 20 25 30
 Thr Asp Ile Leu Ile Asn Leu Met Gln Pro Arg Gln Leu Val Phe Arg
 35 40 45
 Pro Glu Val Leu Trp Asn His Pro Ile Gln Arg Val Ile His Asn Glu
 50 55 60
 Leu Glu Gln Tyr Cys Arg Ala Arg Ala Gly Arg Cys Leu Glu Val Gly
 65 70 75 80
 Ala His Pro Arg Ser Ile Asn Asp Asn Pro Asn Val Leu His Arg Cys
 85 90 95
 Phe Leu Arg Pro Val Gly Arg Asp Val Gln Arg Trp Tyr Ser Ala Pro
 100 105 110
 Thr Arg Gly Pro Ala Ala Asn Cys Arg Arg Ser Ala Leu Arg Gly Leu
 115 120 125
 Pro Pro Val Asp Arg Thr Tyr Cys Xaa Asp
 130 135

<210> 211
 <211> 197
 <212> DNA
 <213> Hepatitis E Virus

<220>
 <223> 1440o2.seq

<221> CDS
 <222> (2)...(196)

<223> Xaa = Unknown or Other at position 3
 <223> Xaa = Unknown or Other at position 60
 <223> Xaa = Unknown or Other at positions 62-63

```

<400> 211
g aca gaa ttr att tcg tcg gct gga ggt caa ctg ttc tac tcc cgc ccg      49
Thr Glu Xaa Ile Ser Ser Ala Gly Gly Gln Leu Phe Tyr Ser Arg Pro
  1             5             10             15

ggt gtc tca gcc aat ggc gag ccg act gtt aag tta tac acc tct gtc      97
Val Val Ser Ala Asn Gly Glu Pro Thr Val Lys Leu Tyr Thr Ser Val
          20             25             30

gag aat gca cag cag gat aag ggc att gct ata cca cat gat ata gac      145
Glu Asn Ala Gln Gln Asp Lys Gly Ile Ala Ile Pro His Asp Ile Asp
          35             40             45

tta ggg gat tcc cgt gtg gtt ata caa gat tat gay aac car cay gaa      193
Leu Gly Asp Ser Arg Val Val Ile Gln Asp Tyr Xaa Asn Xaa Xaa Glu
          50             55             60

caa g      197
Gln
65

```

```

<210> 212
<211> 65
<212> PRT
<213> Hepatitis E Virus

<220>
<223> Xaa = Unknown or Other at position 3

<223> Xaa = Unknown or Other at position 60

<223> Xaa = Unknown or Other at positions 62-63

```

```

<400> 212
Thr Glu Xaa Ile Ser Ser Ala Gly Gly Gln Leu Phe Tyr Ser Arg Pro
  1             5             10             15
Val Val Ser Ala Asn Gly Glu Pro Thr Val Lys Leu Tyr Thr Ser Val
          20             25             30
Glu Asn Ala Gln Gln Asp Lys Gly Ile Ala Ile Pro His Asp Ile Asp
          35             40             45
Leu Gly Asp Ser Arg Val Val Ile Gln Asp Tyr Xaa Asn Xaa Xaa Glu
          50             55             60
Gln
65

```

```

<210> 213
<211> 418
<212> DNA
<213> Hepatitis E Virus

<220>
<223> 2015-1.seq

<221> CDS
<222> (3)...(416)

```

<223> Xaa = Unknown or Other at position 2

<223> Xaa = Unknown or Other at position 5

<223> Xaa = Unknown or Other at position 137

<400> 213

```

ct ggc aty act act gcy att gag cag gct gct ctg gct gcg gct aac      47
Gly Xaa Thr Thr Xaa Ile Glu Gln Ala Ala Leu Ala Ala Ala Asn
  1              5              10              15

tct gcc ttg gcg aat gct gtg gtg gtc cgg ccg ttc ctg tcc cgc act      95
Ser Ala Leu Ala Asn Ala Val Val Val Arg Pro Phe Leu Ser Arg Thr
              20              25              30

cag act gat att ctt att aat ttg atg caa ccc cgg caa ctt gta ttc      143
Gln Thr Asp Ile Leu Ile Asn Leu Met Gln Pro Arg Gln Leu Val Phe
              35              40              45

cgc cct gag gtt ttg tgg aac cat ccg atc cag cga gtc ata cat aat      191
Arg Pro Glu Val Leu Trp Asn His Pro Ile Gln Arg Val Ile His Asn
              50              55              60

gag ctg gag cag tat tgc cgt gct cgt gct ggt cgc tgc ctg gag gtt      239
Glu Leu Glu Gln Tyr Cys Arg Ala Arg Ala Gly Arg Cys Leu Glu Val
              65              70              75

ggg gct cat cca aga tct atc aat gac aac cct aat gtt ctg cac cgg      287
Gly Ala His Pro Arg Ser Ile Asn Asp Asn Pro Asn Val Leu His Arg
              80              85              90              95

tgt ttc ctc cgt ccg gtt ggg cga gac gta cag cgt tgg tat tct gcc      335
Cys Phe Leu Arg Pro Val Gly Arg Asp Val Gln Arg Trp Tyr Ser Ala
              100              105              110

cct act cgc gcc ccg gcg gct aat tgc cgc cgt tcc gcg tta cgt ggc      383
Pro Thr Arg Gly Pro Ala Ala Asn Cys Arg Arg Ser Ala Leu Arg Gly
              115              120              125

cta cct cct gtc gac cgc act tac tgt yty gat gg                      418
Leu Pro Pro Val Asp Arg Thr Tyr Cys Xaa Asp
              130              135

```

<210> 214

<211> 138

<212> PRT

<213> Hepatitis E Virus

<220>

<223> Xaa = Unknown or Other at position 2

<223> Xaa = Unknown or Other at position 5

<223> Xaa = Unknown or Other at position 137

<400> 214

Gly Xaa Thr Thr Xaa Ile Glu Gln Ala Ala Leu Ala Ala Ala Asn Ser
 1 5 10 15
 Ala Leu Ala Asn Ala Val Val Val Arg Pro Phe Leu Ser Arg Thr Gln
 20 25 30
 Thr Asp Ile Leu Ile Asn Leu Met Gln Pro Arg Gln Leu Val Phe Arg
 35 40 45
 Pro Glu Val Leu Trp Asn His Pro Ile Gln Arg Val Ile His Asn Glu
 50 55 60
 Leu Glu Gln Tyr Cys Arg Ala Arg Ala Gly Arg Cys Leu Glu Val Gly
 65 70 75 80
 Ala His Pro Arg Ser Ile Asn Asp Asn Pro Asn Val Leu His Arg Cys
 85 90 95
 Phe Leu Arg Pro Val Gly Arg Asp Val Gln Arg Trp Tyr Ser Ala Pro
 100 105 110
 Thr Arg Gly Pro Ala Ala Asn Cys Arg Arg Ser Ala Leu Arg Gly Leu
 115 120 125
 Pro Pro Val Asp Arg Thr Tyr Cys Xaa Asp
 130 135

<210> 215

<211> 197

<212> DNA

<213> Hepatitis E Virus

<220>

<223> 2015o2.seq

<221> CDS

<222> (2)...(196)

<223> Xaa = Unknown or Other at position 3

<223> Xaa = Unknown or Other at position 60

<223> Xaa = Unknown or Other at positions 62-63

<400> 215

g aca gaa ttr att tcg tcg gct gga ggc cag ctc ttc tac tcc cgc cca 49
 Thr Glu Xaa Ile Ser Ser Ala Gly Gly Gln Leu Phe Tyr Ser Arg Pro
 1 5 10 15
 gtc gtc tca gcc aat ggc gag ccg act gtt aaa ttg tat aca tcc gtc 97
 Val Val Ser Ala Asn Gly Glu Pro Thr Val Lys Leu Tyr Thr Ser Val
 20 25 30
 gag aat gcg cag cag gac aag ggc att gcc ata cca cat gat ata gat 145
 Glu Asn Ala Gln Gln Asp Lys Gly Ile Ala Ile Pro His Asp Ile Asp
 35 40 45
 cta gga gat tcc cgc gtg gtt atc cag gat tat gay aac car cay gaa 193
 Leu Gly Asp Ser Arg Val Val Ile Gln Asp Tyr Xaa Asn Xaa Xaa Glu
 50 55 60
 caa g 197
 Gln
 65

```

<210> 216
<211> 65
<212> PRT
<213> Hepatitis E Virus

<220>
<223> Xaa = Unknown or Other at position 3

<223> Xaa = Unknown or Other at position 60

<223> Xaa = Unknown or Other at positions 62-63

<400> 216
Thr Glu Xaa Ile Ser Ser Ala Gly Gly Gln Leu Phe Tyr Ser Arg Pro
 1          5          10          15
Val Val Ser Ala Asn Gly Glu Pro Thr Val Lys Leu Tyr Thr Ser Val
          20          25          30
Glu Asn Ala Gln Gln Asp Lys Gly Ile Ala Ile Pro His Asp Ile Asp
          35          40          45
Leu Gly Asp Ser Arg Val Val Ile Gln Asp Tyr Xaa Asn Xaa Xaa Glu
 50          55          60
Gln
65

<210> 217
<211> 251
<212> DNA
<213> Hepatitis E Virus

<220>
<223> 14404-2.seq

<221> CDS
<222> (3)...(251)
<223> orf2

<223> orf3 from position 1 to position 165

<400> 217
at att cat cca acc aac ccc ttt gcc tcc gac gtc gta tcg caa tcc          47
  Ile His Pro Thr Asn Pro Phe Ala Ser Asp Val Val Ser Gln Ser
  1          5          10          15

ggg gct gga gct cgc cct cga cag ccg gcc cgc ccc ctc ggc tcc tct          95
Gly Ala Gly Ala Arg Pro Arg Gln Pro Ala Arg Pro Leu Gly Ser Ser
          20          25          30

tgg cgt gac cag tcc cag cgc ccc ccc gct gtc ccc cgt cgt cga tct          143
Trp Arg Asp Gln Ser Gln Arg Pro Pro Ala Val Pro Arg Arg Arg Ser
          35          40          45

acc cca act ggg gct gcg ccg cta act gct gtt tca cca gcg cct gat          191
Thr Pro Thr Gly Ala Ala Pro Leu Thr Ala Val Ser Pro Ala Pro Asp
          50          55          60

acg gcc cca gtc cct gat gtt gac tct cgt ggc gct atc ttg cgc cgg          239

```

Thr Ala Pro Val Pro Asp Val Asp Ser Arg Gly Ala Ile Leu Arg Arg
 65 70 75
 cag tat aac cta
 Gln Tyr Asn Leu
 80

251

<210> 218
 <211> 83
 <212> PRT
 <213> Hepatitis E Virus

<400> 218
 Ile His Pro Thr Asn Pro Phe Ala Ser Asp Val Val Ser Gln Ser Gly
 1 5 10 15
 Ala Gly Ala Arg Pro Arg Gln Pro Ala Arg Pro Leu Gly Ser Ser Trp
 20 25 30
 Arg Asp Gln Ser Gln Arg Pro Pro Ala Val Pro Arg Arg Arg Ser Thr
 35 40 45
 Pro Thr Gly Ala Ala Pro Leu Thr Ala Val Ser Pro Ala Pro Asp Thr
 50 55 60
 Ala Pro Val Pro Asp Val Asp Ser Arg Gly Ala Ile Leu Arg Arg Gln
 65 70 75 80
 Tyr Asn Leu

<210> 219
 <211> 55
 <212> PRT
 <213> Hepatitis E Virus

<220>
 <223> 14404-2.seq orf3

<400> 219
 Ile Phe Ile Gln Pro Thr Pro Leu Pro Pro Thr Ser Tyr Arg Asn Pro
 1 5 10 15
 Gly Leu Glu Leu Ala Leu Asp Ser Arg Pro Ala Pro Ser Ala Pro Leu
 20 25 30
 Gly Val Thr Ser Pro Ser Ala Pro Pro Leu Ser Pro Val Val Asp Leu
 35 40 45
 Pro Gln Leu Gly Leu Arg Arg
 50 55

<210> 220
 <211> 251
 <212> DNA
 <213> Hepatits E Virus

<220>
 <223> 20154-2.seq

<221> CDS
 <222> (3)...(251)
 <223> orf2

<223> orf3 from position 1 to position 165

<400> 220

```

at att cat cca acc aac ccc ttt gcc gcc gac gtc gta tca caa ccc      47
Ile His Pro Thr Asn Pro Phe Ala Ala Asp Val Val Ser Gln Pro
  1             5             10             15

ggg gct gga gct cgc cct cga cag ccg ccc cgc ccc ctc ggc tcc tct      95
Gly Ala Gly Ala Arg Pro Arg Gln Pro Pro Arg Pro Leu Gly Ser Ser
                20             25             30

tgg cgt gat cag tcc cag cgc ccc tcc gct gcc ccc cgt cgt cga tct     143
Trp Arg Asp Gln Ser Gln Arg Pro Ser Ala Ala Pro Arg Arg Arg Ser
                35             40             45

acc cca gct ggg gct gcg ccg tta act gct gtt tcc cct gcg ccc gat     191
Thr Pro Ala Gly Ala Ala Pro Leu Thr Ala Val Ser Pro Ala Pro Asp
                50             55             60

acg gcc cca gtc ccc gac gtt gat tcc cgt ggt gcc atc ctg cgc cgg     239
Thr Ala Pro Val Pro Asp Val Asp Ser Arg Gly Ala Ile Leu Arg Arg
                65             70             75

cag tat aac cta
Gln Tyr Asn Leu
  80

```

<210> 221

<211> 83

<212> PRT

<213> Hepatitis E Virus

<400> 221

```

Ile His Pro Thr Asn Pro Phe Ala Ala Asp Val Val Ser Gln Pro Gly
  1             5             10             15
Ala Gly Ala Arg Pro Arg Gln Pro Pro Arg Pro Leu Gly Ser Ser Trp
                20             25             30
Arg Asp Gln Ser Gln Arg Pro Ser Ala Ala Pro Arg Arg Arg Ser Thr
                35             40             45
Pro Ala Gly Ala Ala Pro Leu Thr Ala Val Ser Pro Ala Pro Asp Thr
                50             55             60
Ala Pro Val Pro Asp Val Asp Ser Arg Gly Ala Ile Leu Arg Arg Gln
                65             70             75             80
Tyr Asn Leu

```

<210> 222

<211> 55

<212> PRT

<213> Hepatitis E Virus

<220>

<223> 20154-2.seq orf3

<400> 222

```

Ile Phe Ile Gln Pro Thr Pro Leu Pro Pro Thr Ser Tyr His Asn Pro

```

```

      1           5           10           15
Gly Leu Glu Leu Ala Leu Asp Ser Arg Pro Ala Pro Ser Ala Pro Leu
      20           25           30
Gly Val Ile Ser Pro Ser Ala Pro Pro Leu Pro Pro Val Val Asp Leu
      35           40           45
Pro Gln Leu Gly Leu Arg Arg
      50           55

```

```

<210> 223
<211> 48
<212> PRT
<213> Hepatitis E Virus

```

```

<220>
<223> US-2 3-2e

```

```

      1           5           10           15
Thr Ile Asp Tyr Pro Ala Arg Ala His Thr Phe Asp Asp Phe Cys Pro
      20           25           30
Glu Cys Arg Thr Leu Gly Leu Gln Gly Cys Ala Phe Gln Ser Thr Ile
      35           40           45
Ala Glu Leu Gln Arg Leu Lys Met Lys Val Gly Lys Thr Arg Glu Ser

```

```

<210> 224
<211> 33
<212> PRT
<213> Hepatitis E Virus

```

```

<220>
<223> US-2 4-2

```

```

      1           5           10           15
Asp Ser Arg Pro Ala Pro Leu Val Pro Leu Gly Val Thr Ser Pro Ser
      20           25           30
Ala Pro Pro Leu Pro Pro Val Val Asp Leu Pro Gln Leu Gly Leu Arg
Arg

```

```

<210> 225
<211> 450
<212> DNA
<213> Hepatitis E Virus

```

```

<220>
<223> 5p.pile {hpesvp}

```

```

<400> 225
ggctcctggc atcactactg ctattgagca ggctgctcta gcagcggcca actctgcct      60
ggcgaatgct gtggtagtta ggccttttct ctctcaccag cagattgaga tctcattaa      120
cctaatagcaa cctcgccagc ttgttttccg ccccgagggt ttctggaatc atcccatcca      180
gcgtgtcatc cataacgagc tggagcttta ctgccgcgcc cgctccggcc gctgtcttga      240
aattggcgcc catccccgct caataaatga taatcctaata gtggtccacc gctgcttct      300
ccgccctggt gggcgatgat ttcagcgctg gtatactgct cccactcgcg ggccggctgc      360
taattgcggg cgttccgcgc tgcgcgggct tcccgctgct gaccgcactt actgcctcga      420
cgggttttct ggctgtaact ttcccgccga                                     450

```

<210> 226
 <211> 450
 <212> DNA
 <213> Hepatitis E Virus

<220>
 <223> 5p.pile {hpeuigh}

<400> 226
 ggctcctggc atcactactg ctattgagca ggctgctcta gcagcggcca attctgccct 60
 tgcgaatgct gtggtagtta ggccttttct ctctcaccag cagattgaga tccttattaa 120
 cctaatagcaa cctcgccagc ttgttttccg ccccgagggt ttctggaacc accccatcca 180
 gcgtgtcatc cataatgagc tggagcttta ctgtcgcgcc cgctccggcc gctgccttga 240
 aattgggtgcc caccctcgct caataaaaga caatcctaata gtggtccacc gctgcttcc 300
 ccgccctgcc gggcgtgatg ttcagcgttg gtatactgct cctacccgcg ggccggctgc 360
 taattgccgg gggtccgcac tgcgcgggct ccccgctgct gaccgcactt actgcttoga 420
 cgggttttct ggctgtaact ttcccgccga 450

<210> 227
 <211> 450
 <212> DNA
 <213> Hepatitis E Virus

<220>
 <223> 5p.pile {hpea}

<400> 227
 ggctcctggc atcactactg ctattgagca ggctgctcta gcagcggcca actctgccct 60
 tgcgaatgct gtggtagtta ggccttttct ctctcaccag cagattgaga tccttattaa 120
 cctaatagcaa cctcgccagc ttgttttccg ccccgagggt ttctggaacc atcccatcca 180
 gcgtgttatc cataatgagc tggagcttta ctgtcgcgcc cgctccggcc gctgcctcga 240
 aattgggtgcc caccctcgct caataaatga caatcctaata gtggtccacc gttgcttcc 300
 ccgtcctgcc gggcgtgatg ttcagcgttg gtatactgcc cctacccgcg ggccggctgc 360
 taattgccgg cggtccgcgc tgcgcgggct ccccgctgct gaccgcactt actgcttoga 420
 cgggttttct ggctgtaact ttcccgccga 450

<210> 228
 <211> 446
 <212> DNA
 <213> Hepatitis E Virus

<220>
 <223> 5p.pile {840455p}

<400> 228
 cctggcatta ctactgcoat tgagcaggct gctctggctg cggccaattc tgccttggcg 60
 aatgctgtgg tgggtcggcc gtttttatct cgcgtgcaaa ccgagattct tattaatttg 120
 atgcaacccc ggcagttggg ttcccgccct gaggtacttt ggaatcacc tatccagcgg 180
 gttatacata atgaattaga acagtactgc cgggctcggg ctggtcgttg cttggagggt 240
 ggagctcacc caagatccat taatgacaac cccaacgttc tgcacgggtg tttccttaga 300
 ccggttggcc gagatgttca gcgctggtac tctgccccca ccccgggccc tgcggctaata 360
 tgccgcgcgt ccgcgttgcg tggctctccc cccgctgacc gcacttactg ctttgatgga 420
 ttctccggtt gtgcttttgc tgcaga 446

<210> 229
 <211> 450

<212> DNA
<213> Hepatitis E Virus

<220>
<223> 5p.pile {hpenssp}

```

<400> 229
ggctcctggc atcactactg ctattgagca agcagctcta gcagcggcca actccgccct    60
tgcaaatgct gtgggtggcc ggcctttcct ttoccatcag cagggttgaga tccttataaa    120
tctcatgcaa cctcggcagc tgggtgttcg tcctgaggtt ttttggaatc acccgattca    180
acgtgttata cataatgagc ttgagcagta ttgccgtgct cgctcgggtc gctgccttga    240
gattggagcc caccacagct ccattaatga taatcctaata gtccctccatc gctgctttct    300
ccaccccgtc ggcggggatg ttcagcgctg gtacacagcc ccgactaggg gacctgcggc    360
gaactgtcgc cgctcggcac ttcgtggtct gccaccagcc gaccgcactt actgttttga    420
tggctttgcc ggctgccggt ttgccgccga

```

<210> 230
<211> 450
<212> DNA
<213> Hepatitis E Virus

<220>
<223> 5p Consensus

<221> variation
<222> (1)...(450)
<223> n = a or g or c or t/u, unknown or other in each instance and is indicated in Figure 9

```

<400> 230
nnnnccctggc atnactactg cnattgagca ngcngctctn gcngcggcca antcngccnt    60
ngcgaatgct gtggtngtnn ggcctttnt ntncnnnnng cannnngaga tncnnaatnaa    120
nntnatgcaa ccncgncagn tngtnttncg nccngaggtt ntntggaanc anccnatnca    180
ncngtgnatn cataangann tngancnnta ntgncngncn cgnnncngnc gntgntntga    240
nnttggngcn canccnngnt cnatnaanga naancnnaan gtntntcanc gntgntntct    300
nnnnccngnn ggcngngatg ttcagcgntg gtanncngcn ccnacnngng gncncngncg    360
naantgncgn ngntcngcnn tncngngnct nccnnncngcn gaccgcactt actgntntga    420
nngntntnncn ngntgnnnnnt ttncngcnga

```

<210> 231
<211> 300
<212> DNA
<213> Hepatitis E Virus

<220>
<223> 3p.pile {hpea} shown in Figure 9B

```

<400> 231
actgagtcag tgaagccagt gcttgacctg acaaattcaa ttctgtgtcg ggtggaatga    60
ataacatgct ttttgctgcg cccatgggtt cgcgaccatg cgccctcggc ctattttgct    120
gttgctcctc atgtttctgc ctatgtgccc cgcgccaccg cccggtcagc cgtctggccg    180
cgctcgtggg cggcgagcgc gcggttccgg cgggtggttc tggggtgacc ggggtgattc    240
tcagcccttc gcaatccctt atattcatcc aaccaacccc ttccgccccg atgtcaccgc    300

```

<210> 232
<211> 300
<212> DNA

<213> Hepatitis E Virus

<220>

<223> 3p.pile {hpeuigh} shown in Figure 9B

<400> 232

```
actgagtcgg tgaagccagt gctcgacttg acaaattcaa tcctgtgtcg ggtggaatga    60
ataacatgtc ttttgctgcg cccatgggtt ggcgaccatg cgcctcggc ctattttgt    120
gttgctcttc atgtttctgc ctatcgtgcc cgcgccaccg cccggtcagc cgtctggccg    180
ccgtcgtggg cggcgccagcg gcggttcggg cgggtggttc tggggtgacc gggttgattc    240
tcagcccttc gcaatccct atattcatcc aaccaacccc ttgcctcccg atgtcacgcg    300
```

<210> 233

<211> 300

<212> DNA

<213> Hepatitis E Virus

<220>

<223> 3p.pile {hpesvp} shown in Figure 9B

<400> 233

```
actgagtcag taaaaccagt gctcgacttg acaaattcaa tcttgtgtcg ggtggaatga    60
ataacatgtc ttttgctgcg cccatgggtt cgcgaccatg cgcctcggc ctattttgtt    120
gctgctcttc atgttttctgc ctatcgtgcc cgcgccaccg cccggtcagc cgtctggccg    180
ccgtcgtggg cggcgccagcg gcggttcggg cgggtggttc tggggtgacc gggttgattc    240
tcagcccttc gcaatccct atattcatcc aaccaacccc ttgcctcccg atgtcacgcg    300
```

<210> 234

<211> 300

<212> DNA

<213> Hepatitis E Virus

<220>

<223> 3p.pile {hpenssp} shown in Figure 9B

<400> 234

```
acagagtcctg ttaagcctat acttgacctt acacactcaa ttatgcaccg gtctgaatga    60
ataacatgtg gtttgctgcg cccatgggtt cgcgaccatg cgccttaggc ctcttttgtt    120
gttggtcttc ttgtttctgc ctatgttgcc cgcgccaccg accggtcagc cgtctggccg    180
ccgtcgtggg cggcgccagcg gcggtaccgg cgggtggttc tggggtgacc gggttgattc    240
tcagcccttc gcaatccct atattcatcc aaccaacccc ttgccccag acgttgccgc    300
```

<210> 235

<211> 297

<212> DNA

<213> Hepatitis E Virus

<220>

<223> 3p.pile {840453p} shown in Figure 9B

<400> 235

```
acagagacta ttaaacctgt acttgatctc acaaattcca tcatacagcg ggtggaatga    60
ataacatgtc ttttgcatcg cccatgggat caccatgcgc cctagggctg ttctgttggt    120
gttctctcatg tttctgccta tgcgtcccg gccaccggcc ggtcagccgt ctggccgctcg    180
ccgtgggcgg cgcagcgcg gtgcggcggt tgggttcttg agtgacaggg ttgattctca    240
gcccttcgcc ctcccctata ttcatccaac caacccttc gccgccgatg tegtctc    297
```

<210> 236
 <211> 300
 <212> DNA
 <213> Hepatitis E Virus

<220>
 <223> 3p Consensus shown in Figure 9B

<221> variation
 <222> (3)...(300)
 <223> n = a or g or c or t/u, unknown or other in each instance and is indicated in Figure 9B

```

<400> 236
acngagncnn tnaancnnt nctnganntn acanantcna tnntnnnnncg gnnngaataga      60
ataacatgtn ntttgcnncg cccatgggnt nnnnaccatg cgccctnggn ctnttntgnt      120
gntgntoctc ntgtttntgc ctatnmtgcc cgcgccaccg nccggtcagc cgtctggccg      180
ncgncgtggg cggcgcagcg gcggtncggg cgggtggttc tggngtgacn gggttgattc      240
tcagcccttc gcnntcccct atattcatcc aaccaacccc ttngccncng angtnnnnnnc      300

```

<210> 237
 <211> 250
 <212> DNA
 <213> Hepatitis E Virus

<220>
 <223> 3p.pile {hpea} shown in Figure 9C

```

<400> 237
agcgcttacc ctgtttaacc ttgctgacac cctgcttggc ggtctaccga cagaattgat      60
ttcgctgggt ggtggccagc tgttctactc tcgccccgtc gtctcagcca atggcgagcc      120
gactgttaag ctgtatacat ctgtggagaa tgctcagcag gataagggtg ttgcaatccc      180
gcatgacatc gacctcgggg aatcccgtgt agttattcag gattatgaca accaacaatga      240
gcaggaccga                                     250

```

<210> 238
 <211> 250
 <212> DNA
 <213> Hepatitis E Virus

<220>
 <223> 3p.pile {hpeuigh} shown in Figure 9C

```

<400> 238
agcgcttacc ctgtttaacc ttgctgacac cctgcttggc ggtctaccga cagaattgat      60
ttcgctgggt ggtggccagc tgttctactc tcgccccgtc gtctcagcca atggcgagcc      120
gactgttaag ctgtatacat ctgtagagaa tgctcagcag gataagggtg ttgcaatccc      180
gcatgacatc gacctcgggg aatctcgagt tgttattcag gattatgaca accaacaatga      240
gcaggaccgg                                     250

```

<210> 239
 <211> 250
 <212> DNA
 <213> Hepatitis E Virus

<220>
 <223> 3p.pile {hpesvp} shown in Figure 9C

```

<400> 239
agccctcacc ctgttcaacc ttgctgacac tctgcttgge ggcctgccga cagaattgat    60
ttcgtcggct ggtggccagc tgttctactc ccgtcccggt gtctcagcca atggcgagcc    120
gactgttaag ttgtatacat ctgtagagaa tgctcagcag gataagggtg ttgcaatccc    180
gcatgacatt gacctcggag aatctcgtgt ggttattcag gattatgata accaacatga    240
acaagatcgg                                     250

```

```

<210> 240
<211> 250
<212> DNA
<213> Hepatitis E Virus

```

```

<220>
<223> 3p.pile {hpenssp} shown in Figure 9C

```

```

<400> 240
agctctaaca ttacttaacc ttgctgacac gctcctcgge gggctcccgga cagaattaat    60
ttcgtcggct ggcgggcaac tgttttattc ccgcccggtt gtctcagcca atggcgagcc    120
aacgtgaag ctctatacat cagtggagaa tgctcagcag gataagggtg ttgctatccc    180
ccacgatatc gatcttgggtg attcgcgtgt ggtcattcag gattatgata accagcatga    240
gcaggatcgg                                     250

```

```

<210> 241
<211> 250
<212> DNA
<213> Hepatitis E Virus

```

```

<220>
<223> 3p.pile {840453p} shown in Figure 9C

```

```

<400> 241
tgccctgact ctgtttaac ttgctgatac gcttcttggt ggtttaccga cagaattgat    60
ttcgtcggct ggggggcaac tgttttactc ccgcccggtt cagaattgat ttcgtcggct    120
gggggtcaac tgttttactc ccgccctggt tgccgcagcaa gacaagggtg tcaccattcc    180
acacgacata gatttaggtg actcccggtt ggttatccag gattatgata accagcacga    240
acaagatcga                                     250

```

```

<210> 242
<211> 250
<212> DNA
<213> Hepatitis E Virus

```

```

<220>
<223> 3p Consensus shown in Figure 9C

```

```

<221> variation
<222> (1)...(250)
<223> n = a or g or c or t/u, unknown or other at each
       instance and is indicated in Figure 9C

```

```

<400> 242
ngcnctnacn ntntnaanc ttgctganac nctnctnggn ggnntnccga cagaattnat    60
ttcgtcggct gngngncaac tgttntantc ccgncngtn gtctcngcca atggcgagcc    120
nacngtinaag ntntanacat cngtngagaa tgcncagcan ganaagggnn tnnctatncc    180
ncanganatn ganntnggng antcncnggt ngtnatncag gattatgana accancanga    240
ncanganacn                                     250

```

```

<210> 243
<211> 418
<212> DNA
<213> Hepatitis E Virus

<220>
<223> Aulo1-wlabolp1.pat

<221> CDS
<222> (3)...(416)

<223> Xaa = Unknown or Other at position 2

<223> Xaa = Unknown or Other at position 5

<223> Xaa = Unknown or Other at position 137

<400> 243
ct ggc aty act act gcy att gag caa gct gct ctg gct gcg gcc aat      47
  Gly Xaa Thr Thr Xaa Ile Glu Gln Ala Ala Leu Ala Ala Ala Asn
    1              5              10              15

tct gcc ttg gcg aat gct gtg gtg gtt cgg ccg ttt tta tcc cgt gtg      95
  Ser Ala Leu Ala Asn Ala Val Val Val Arg Pro Phe Leu Ser Arg Val
                20              25              30

cag act gag atc ctt att aac ttg atg caa cct cgg cag ctg gtg ttc      143
  Gln Thr Glu Ile Leu Ile Asn Leu Met Gln Pro Arg Gln Leu Val Phe
                35              40              45

cga cct gag gtg ctt tgg aat cat ccc att cag cgg gtt atc cat aat      191
  Arg Pro Glu Val Leu Trp Asn His Pro Ile Gln Arg Val Ile His Asn
                50              55              60

gag tta gaa caa tac tgc cgg gcc cgg gcc gcc cgt tgc cta gag gtg      239
  Glu Leu Glu Gln Tyr Cys Arg Ala Arg Ala Gly Arg Cys Leu Glu Val
                65              70              75

ggg gcc cac cca agg tcc att aac gat aac ccc aat gtt ttg cac cgg      287
  Gly Ala His Pro Arg Ser Ile Asn Asp Asn Pro Asn Val Leu His Arg
                80              85              90              95

tgt ttt ctg cga ccg gtc ggg agg gat gtt cag cgc tgg tac tot gcc      335
  Cys Phe Leu Arg Pro Val Gly Arg Asp Val Gln Arg Trp Tyr Ser Ala
                100              105              110

ccc acc cgc gcc cct gcg gct aac tgc cgc cgc tcc gct ttg cgt gcc      383
  Pro Thr Arg Gly Pro Ala Ala Asn Cys Arg Arg Ser Ala Leu Arg Gly
                115              120              125

ctt ccc ccc gtc gac cgc act tac tgt yty gat gg                      418
  Leu Pro Pro Val Asp Arg Thr Tyr Cys Xaa Asp
                130              135

```

<210> 244

<211> 138
 <212> PRT
 <213> Hepatitis E Virus

<220>
 <223> Xaa = Unknown or Other at position 2
 <223> Xaa = Unknown or Other at position 5
 <223> Xaa = Unknown or Other at position 137

<400> 244
 Gly Xaa Thr Thr Xaa Ile Glu Gln Ala Ala Leu Ala Ala Asn Ser
 1 5 10 15
 Ala Leu Ala Asn Ala Val Val Val Arg Pro Phe Leu Ser Arg Val Gln
 20 25 30
 Thr Glu Ile Leu Ile Asn Leu Met Gln Pro Arg Gln Leu Val Phe Arg
 35 40 45
 Pro Glu Val Leu Trp Asn His Pro Ile Gln Arg Val Ile His Asn Glu
 50 55 60
 Leu Glu Gln Tyr Cys Arg Ala Arg Ala Gly Arg Cys Leu Glu Val Gly
 65 70 75 80
 Ala His Pro Arg Ser Ile Asn Asp Asn Pro Asn Val Leu His Arg Cys
 85 90 95
 Phe Leu Arg Pro Val Gly Arg Asp Val Gln Arg Trp Tyr Ser Ala Pro
 100 105 110
 Thr Arg Gly Pro Ala Ala Asn Cys Arg Arg Ser Ala Leu Arg Gly Leu
 115 120 125
 Pro Pro Val Asp Arg Thr Tyr Cys Xaa Asp
 130 135

<210> 245
 <211> 197
 <212> DNA
 <213> Hepatitis E Virus

<220>
 <223> Aulo2-wlao2.pat
 <221> CDS
 <222> (2)...(196)
 <223> Xaa = Unknown or Other at position 3
 <223> Xaa = Unknown or Other at position 17
 <223> Xaa = Unknown or Other at position 60
 <223> Xaa = Unknown or Other at positions 62-63

<400> 245
 g aca gaa ttr att tcg tcg gct ggg gga cag tta ttc tac tcc cgc cct 49
 Thr Glu Xaa Ile Ser Ser Ala Gly Gly Gln Leu Phe Tyr Ser Arg Pro
 1 5 10 15
 gty gtc tca gcc aat ggc gag ccg act gtt aaa tta tat aca tct gta 97
 Xaa Val Ser Ala Asn Gly Glu Pro Thr Val Lys Leu Tyr Thr Ser Val

```

                20                25                30
gag aat gcg cag cag gac aag ggg att gcc atc cca cat gat ata gat      145
Glu Asn Ala Gln Gln Asp Lys Gly Ile Ala Ile Pro His Asp Ile Asp
                35                40                45

ctg ggc gac tct cgt gtg gtg atc cag gat tat gay aac car cay gaa      193
Leu Gly Asp Ser Arg Val Val Ile Gln Asp Tyr Xaa Asn Xaa Xaa Glu
                50                55                60

caa g      197
Gln
65

```

```

<210> 246
<211> 65
<212> PRT
<213> Hepatitis E Virus

```

```

<220>
<223> Xaa = Unknown or Other at position 3

<223> Xaa = Unknown or Other at position 17

<223> xaa = Unknown or Other at position 60

<223> Xaa = Unknown or Other at positions 62-63

```

```

<400> 246
Thr Glu Xaa Ile Ser Ser Ala Gly Gly Gln Leu Phe Tyr Ser Arg Pro
 1          5          10          15
Xaa Val Ser Ala Asn Gly Glu Pro Thr Val Lys Leu Tyr Thr Ser Val
                20                25                30
Glu Asn Ala Gln Gln Asp Lys Gly Ile Ala Ile Pro His Asp Ile Asp
                35                40                45
Leu Gly Asp Ser Arg Val Val Ile Gln Asp Tyr Xaa Asn Xaa Xaa Glu
                50                55                60
Gln
65

```

```

<210> 247
<211> 418
<212> DNA
<213> Hepatitis E Virus

```

```

<220>
<223> Ar101-f73olpl.pat

```

```

<221> CDS
<222> (3)...(416)

```

```

<223> Xaa = Unknown or Other at position 2

<223> Xaa = Unknown or Other at position 5

<223> Xaa = Unknown or Other at position 137

```

<400> 247
 ct ggc aty act act gcy att gag caa gct gct ctg gct gcg gcc aac 47
 Gly Xaa Thr Thr Xaa Ile Glu Gln Ala Ala Leu Ala Ala Ala Asn
 1 5 10 15
 tct gcc ttg gcg aat gct gtg gtg gtt cgg ccg ttt tta tcc cgt gtg 95
 Ser Ala Leu Ala Asn Ala Val Val Val Arg Pro Phe Leu Ser Arg Val
 20 25 30
 cag acc gag att ctt att aac cta atg caa ccc cgg cag ctg gtt ttt 143
 Gln Thr Glu Ile Leu Ile Asn Leu Met Gln Pro Arg Gln Leu Val Phe
 35 40 45
 cgt cct gag gtg ctt tgg aac cat cct atc cag cgg gtt att cat aat 191
 Arg Pro Glu Val Leu Trp Asn His Pro Ile Gln Arg Val Ile His Asn
 50 55 60
 gag tta gaa cag tac tgt cgg gct cgg gct ggt cgc tgc cta gag gtc 239
 Glu Leu Glu Gln Tyr Cys Arg Ala Arg Ala Gly Arg Cys Leu Glu Val
 65 70 75
 ggg gcc cac cca agg tcc att aat gat aac cct aat gtt ttg cac cgg 287
 Gly Ala His Pro Arg Ser Ile Asn Asp Asn Pro Asn Val Leu His Arg
 80 85 90 95
 tgc ttc cta cga cca gtc ggg agg gat gtt caa cgt tgg tat tcc gcc 335
 Cys Phe Leu Arg Pro Val Gly Arg Asp Val Gln Arg Trp Tyr Ser Ala
 100 105 110
 ccc acc cgc ggt cct gct gcc aac tgc cgc cgt tcc gct ctg cgc gcc 383
 Pro Thr Arg Gly Pro Ala Ala Asn Cys Arg Arg Ser Ala Leu Arg Gly
 115 120 125
 ctc cct ccc gtc gac cgc act tac tgt yty gat gg 418
 Leu Pro Pro Val Asp Arg Thr Tyr Cys Xaa Asp
 130 135

<210> 248
 <211> 138
 <212> PRT
 <213> Hepatitis E Virus

<220>
 <223> Xaa = Unknown or Other at position 2
 <223> Xaa = Unknown or Other at position 5
 <223> Xaa = Unknown or Other at position 137

<400> 248
 Gly Xaa Thr Thr Xaa Ile Glu Gln Ala Ala Leu Ala Ala Ala Asn Ser
 1 5 10 15
 Ala Leu Ala Asn Ala Val Val Val Arg Pro Phe Leu Ser Arg Val Gln
 20 25 30
 Thr Glu Ile Leu Ile Asn Leu Met Gln Pro Arg Gln Leu Val Phe Arg

```

      35      40      45
Pro Glu Val Leu Trp Asn His Pro Ile Gln Arg Val Ile His Asn Glu
  50      55      60
Leu Glu Gln Tyr Cys Arg Ala Arg Ala Gly Arg Cys Leu Glu Val Gly
  65      70      75      80
Ala His Pro Arg Ser Ile Asn Asp Asn Pro Asn Val Leu His Arg Cys
      85      90      95
Phe Leu Arg Pro Val Gly Arg Asp Val Gln Arg Trp Tyr Ser Ala Pro
      100      105      110
Thr Arg Gly Pro Ala Ala Asn Cys Arg Arg Ser Ala Leu Arg Gly Leu
      115      120      125
Pro Pro Val Asp Arg Thr Tyr Cys Xaa Asp
      130      135

```

<210> 249
 <211> 145
 <212> DNA
 <213> Hepatitis E Virus

<220>
 <223> Ar1-f73o2p2.pat

<221> CDS
 <222> (1)...(144)

<223> Xaa = Unknown or Other at position 1

<223> Xaa = Unknown or Other at position 3

<223> Xaa = Unknown or Other at position 44

<223> Xaa = Unknown or Other at positions 46-47

```

      <400> 249
gtg gtc tcr gcc aat ggc gag ccg act gtt aag cta tat aca tct gta      48
Xaa Val Xaa Ala Asn Gly Glu Pro Thr Val Lys Leu Tyr Thr Ser Val
  1      5      10      15

gag aac gcg cag cag gat aaa ggg atc gcc att cca cac gat ata gat      96
Glu Asn Ala Gln Gln Asp Lys Gly Ile Ala Ile Pro His Asp Ile Asp
      20      25      30

ctg ggc gat tcc cgt gtg gtc att cag gat tat gay aac car cay gaa      144
Leu Gly Asp Ser Arg Val Val Ile Gln Asp Tyr Xaa Asn Xaa Xaa Glu
      35      40      45

c      145

```

<210> 250
 <211> 48
 <212> PRT
 <213> Hepatitis E Virus

<220>
 <223> Xaa = Unknown or Other at position 1
 <223> Xaa = Unknown or Other at position 3

<223> Xaa = Unknown or Other at position 44

<223> Xaa = Unknown or Other at positions 46-47

<400> 250

Xaa	Val	Xaa	Ala	Asn	Gly	Glu	Pro	Thr	Val	Lys	Leu	Tyr	Thr	Ser	Val
1				5					10					15	
Glu	Asn	Ala	Gln	Gln	Asp	Lys	Gly	Ile	Ala	Ile	Pro	His	Asp	Ile	Asp
			20					25					30		
Leu	Gly	Asp	Ser	Arg	Val	Val	Ile	Gln	Asp	Tyr	Xaa	Asn	Xaa	Xaa	Glu
		35					40					45			

<210> 251

<211> 418

<212> DNA

<213> Hepatitis E Virus

<220>

<223> Ar2ol-f77olp1.pat

<221> CDS

<222> (3)...(416)

<223> Xaa = Unknown or Other at position 2

<223> Xaa = Unknown or Other at position 5

<223> Xaa = Unknown or Other at position 41

<223> Xaa = Unknown or Other at position 44

<223> Xaa = Unknown or Other at position 93

<223> Xaa = Unknown or Other at position 137

<400> 251

ct	ggc	aty	act	act	gcy	att	gag	caa	gct	gct	ctg	gct	gcg	gct	aac	47
	Gly	Xaa	Thr	Thr	Xaa	Ile	Glu	Gln	Ala	Ala	Leu	Ala	Ala	Ala	Asn	
1					5				10					15		
tct	gcc	ttg	gcg	aat	gct	gtg	gtg	ggt	cgg	cgg	ttt	cta	tcc	cgt	gtg	95
Ser	Ala	Leu	Ala	Asn	Ala	Val	Val	Val	Arg	Pro	Phe	Leu	Ser	Arg	Val	
			20					25					30			
cag	act	gag	atc	ctt	att	aac	tta	atg	car	ccc	cgg	car	ctg	ggt	ttc	143
Gln	Thr	Glu	Ile	Leu	Ile	Asn	Leu	Met	Xaa	Pro	Arg	Xaa	Leu	Val	Phe	
			35				40					45				
cgt	ccc	gag	gtg	ctt	tgg	aat	cat	ccc	att	caa	cgg	ggt	att	cat	aat	191
Arg	Pro	Glu	Val	Leu	Trp	Asn	His	Pro	Ile	Gln	Arg	Val	Ile	His	Asn	
		50				55					60					
gaa	tta	gag	cag	tac	tgc	cgg	acc	cgg	gct	ggc	cgt	tgt	tta	gag	gtc	239
Glu	Leu	Glu	Gln	Tyr	Cys	Arg	Thr	Arg	Ala	Gly	Arg	Cys	Leu	Glu	Val	
		65				70					75					

```

gga gcc cat cca agg tcc att aat gac aac cct aac gtt cyg cac cgg      287
Gly Ala His Pro Arg Ser Ile Asn Asp Asn Pro Asn Val Xaa His Arg
80          85          90          95

tgc ttc tta cga cca gtc ggg agg gat gtc caa cga tgg tac tca gcc      335
Cys Phe Leu Arg Pro Val Gly Arg Asp Val Gln Arg Trp Tyr Ser Ala
          100          105          110

ccc act cgc ggc cct gcg gct aat tgc cgt cgt tcc gct ttg cgt ggt      383
Pro Thr Arg Gly Pro Ala Ala Asn Cys Arg Arg Ser Ala Leu Arg Gly
          115          120          125

ctc cct cct gtc gac cgc act tac tgt yty gat gg                        418
Leu Pro Pro Val Asp Arg Thr Tyr Cys Xaa Asp
          130          135

```

<210> 252
 <211> 138
 <212> PRT
 <213> Hepatitis E Virus

<220>
 <223> Xaa = Unknown or Other at position 2
 <223> Xaa = Unknown or Other at position 5
 <223> Xaa = Unknown or Other at position 41
 <223> Xaa = Unknown or Other at position 44
 <223> Xaa = Unknown or Other at position 93
 <223> Xaa = Unknown or Other at position 137

```

<400> 252
Gly Xaa Thr Thr Xaa Ile Glu Gln Ala Ala Leu Ala Ala Ala Asn Ser
1          5          10          15
Ala Leu Ala Asn Ala Val Val Val Arg Pro Phe Leu Ser Arg Val Gln
20          25          30
Thr Glu Ile Leu Ile Asn Leu Met Xaa Pro Arg Xaa Leu Val Phe Arg
35          40          45
Pro Glu Val Leu Trp Asn His Pro Ile Gln Arg Val Ile His Asn Glu
50          55          60
Leu Glu Gln Tyr Cys Arg Thr Arg Ala Gly Arg Cys Leu Glu Val Gly
65          70          75          80
Ala His Pro Arg Ser Ile Asn Asp Asn Pro Asn Val Xaa His Arg Cys
85          90          95
Phe Leu Arg Pro Val Gly Arg Asp Val Gln Arg Trp Tyr Ser Ala Pro
100          105          110
Thr Arg Gly Pro Ala Ala Asn Cys Arg Arg Ser Ala Leu Arg Gly Leu
115          120          125
Pro Pro Val Asp Arg Thr Tyr Cys Xaa Asp
130          135

```

<210> 253
 <211> 197

```

<212> DNA
<213> Hepatitis E Virus

<220>
<223> Ar2o2-f7702.pat

<221> CDS
<222> (2)...(196)

<223> Xaa = Unknown or Other at position 3

<223> Xaa = Unknown or Other at position 60

<223> Xaa = Unknown or Other at positions 62-63

<400> 253
g aca gaa ttr att tcg tcg gct ggg ggt cag ttg ttt tac tcc cgc cct      49
Thr Glu Xaa Ile Ser Ser Ala Gly Gly Gln Leu Phe Tyr Ser Arg Pro
  1             5             10             15

gtc gtc tca gcc aat ggc gag ccg act gtt aag ttg tat aca tct gtg      97
Val Val Ser Ala Asn Gly Glu Pro Thr Val Lys Leu Tyr Thr Ser Val
      20             25             30

gag aat gcg cag cag gat aaa gga atc gcc atc cca cac gac ata gat      145
Glu Asn Ala Gln Gln Asp Lys Gly Ile Ala Ile Pro His Asp Ile Asp
      35             40             45

ctg ggc gat tcc cgt gtg gtt att cag gat tat gay aac car cay gaa      193
Leu Gly Asp Ser Arg Val Val Ile Gln Asp Tyr Xaa Asn Xaa Xaa Glu
      50             55             60

caa g      197
Gln
65

<210> 254
<211> 65
<212> PRT
<213> Hepatitis E Virus

<220>
<223> Xaa = Unknown or Other at position 3

<223> Xaa = Unknown or Other at position 60

<223> Xaa = Unknown or Other at positions 62-63

<400> 254
Thr Glu Xaa Ile Ser Ser Ala Gly Gly Gln Leu Phe Tyr Ser Arg Pro
  1             5             10             15
Val Val Ser Ala Asn Gly Glu Pro Thr Val Lys Leu Tyr Thr Ser Val
      20             25             30
Glu Asn Ala Gln Gln Asp Lys Gly Ile Ala Ile Pro His Asp Ile Asp
      35             40             45
Leu Gly Asp Ser Arg Val Val Ile Gln Asp Tyr Xaa Asn Xaa Xaa Glu

```

	50	55	60	
Gln				
65				
	<210> 255			
	<211> 23			
	<212> DNA			
	<213> Hepatits E Virus			
	<220>			
	<223> HEVConsORF 1N-a1			
	<400> 255			
certcrarrc	artaggtgcg	gtc		23
	<210> 256			
	<211> 25			
	<212> DNA			
	<213> Hepatits E Virus			
	<220>			
	<223> HEVConsORF 2N-a1			
	<400> 256			
cytgytcttg	ytggttrtca	taatc		25
	<210> 257			
	<211> 21			
	<212> DNA			
	<213> Hepatits E Virus			
	<220>			
	<223> HEVConsORF 1N-s2			
	<400> 257			
cygccytkgc	gaatgctgtg	g		21
	<210> 258			
	<211> 25			
	<212> DNA			
	<213> Hepatits E Virus			
	<220>			
	<223> HEVConsORF 2N-a2			
	<400> 258			
gytctrgygtg	rttrtcataa	tcctg		25

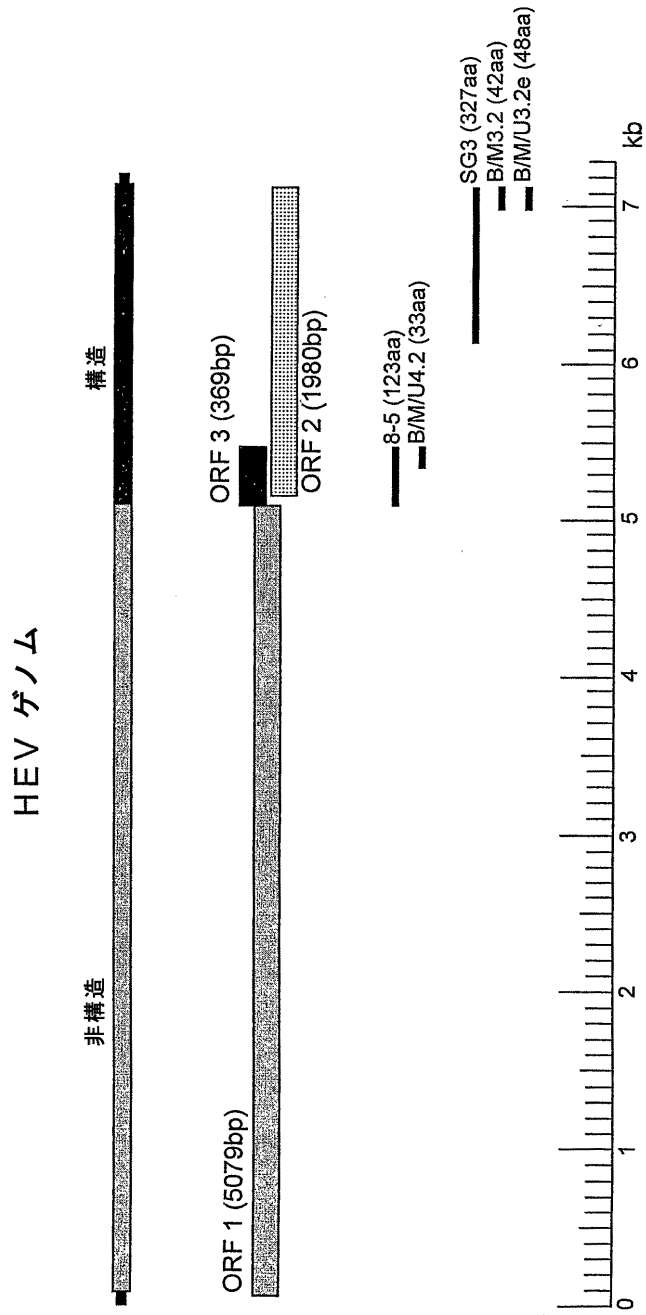


Figure 1

【图 2】

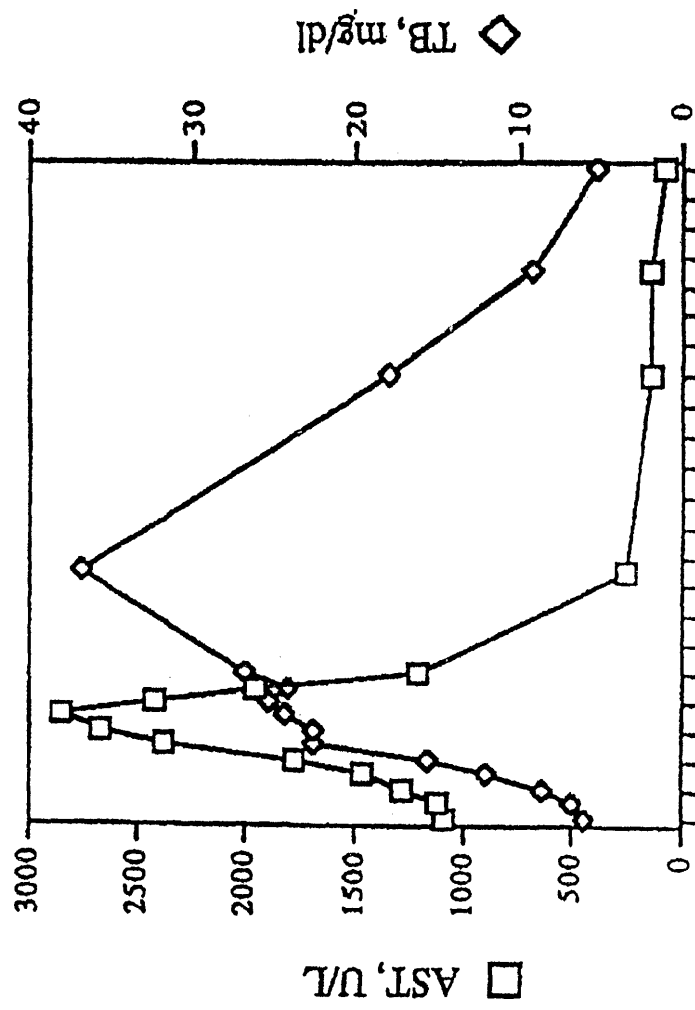


Figure 2

HEV US-1 ゲノム

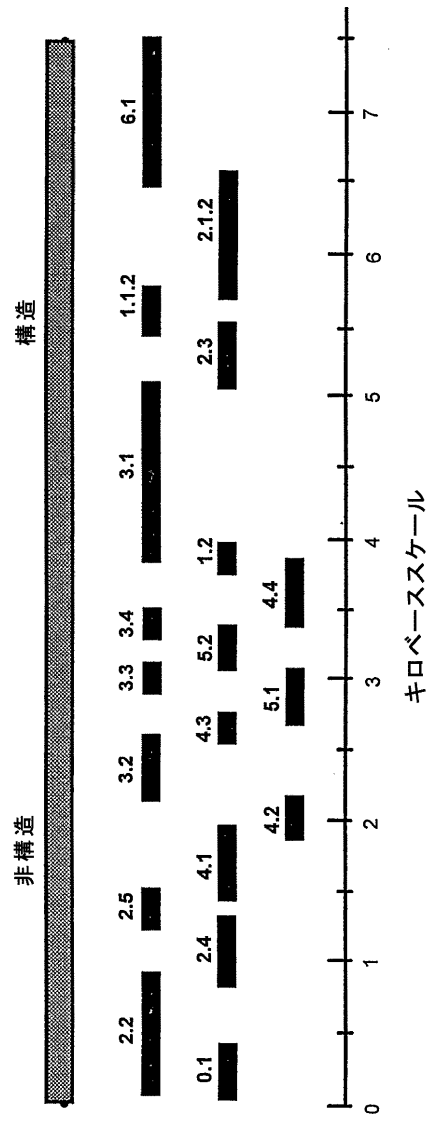


Figure 3

HEV US-2

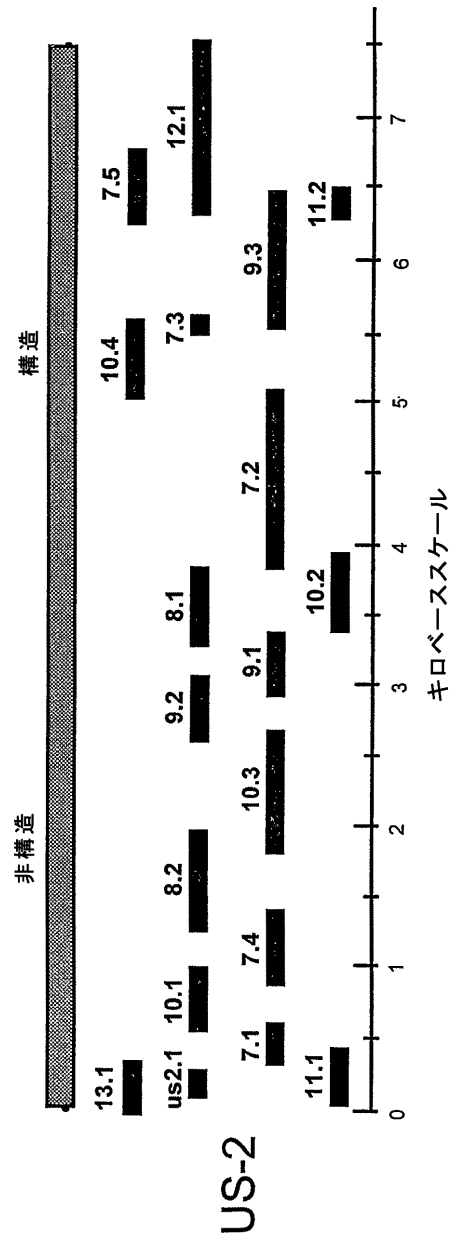
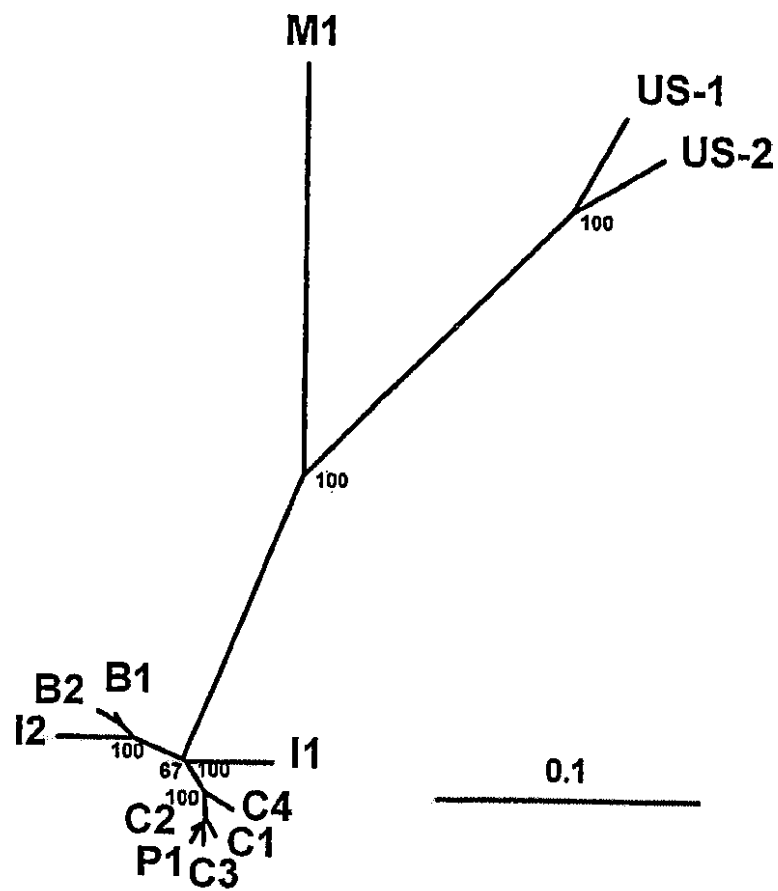
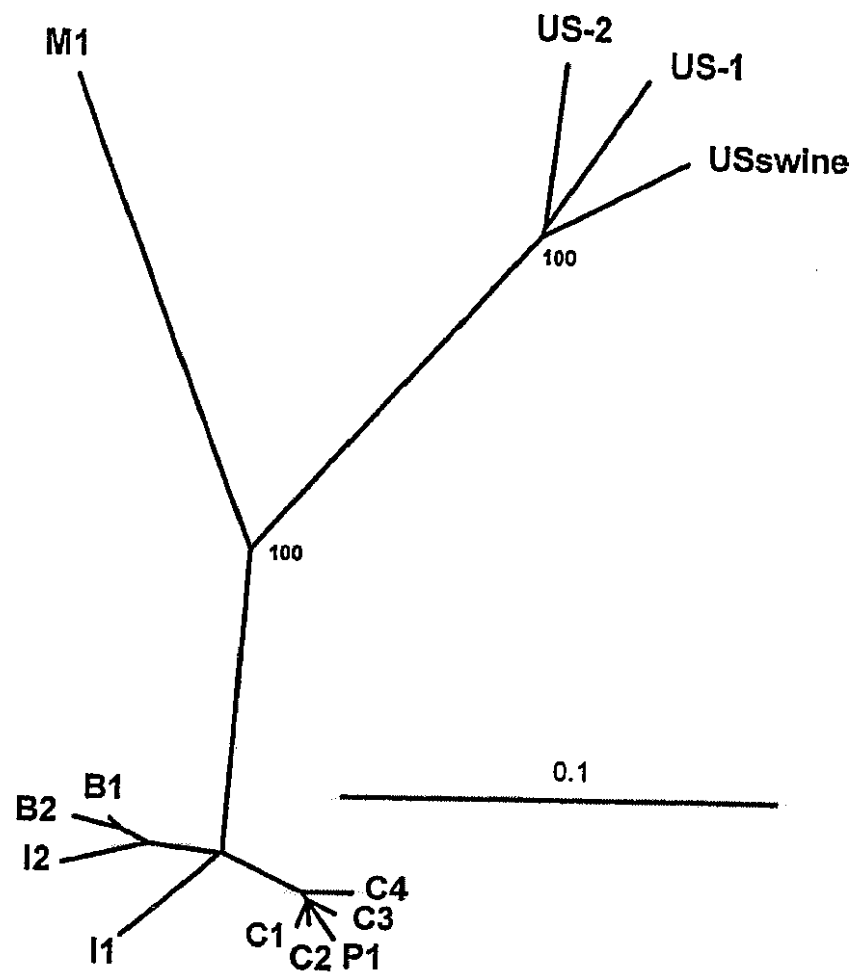


Figure 4

【図5】

*Figure 5*

【図 6】

*Figure 6*

【図 7】

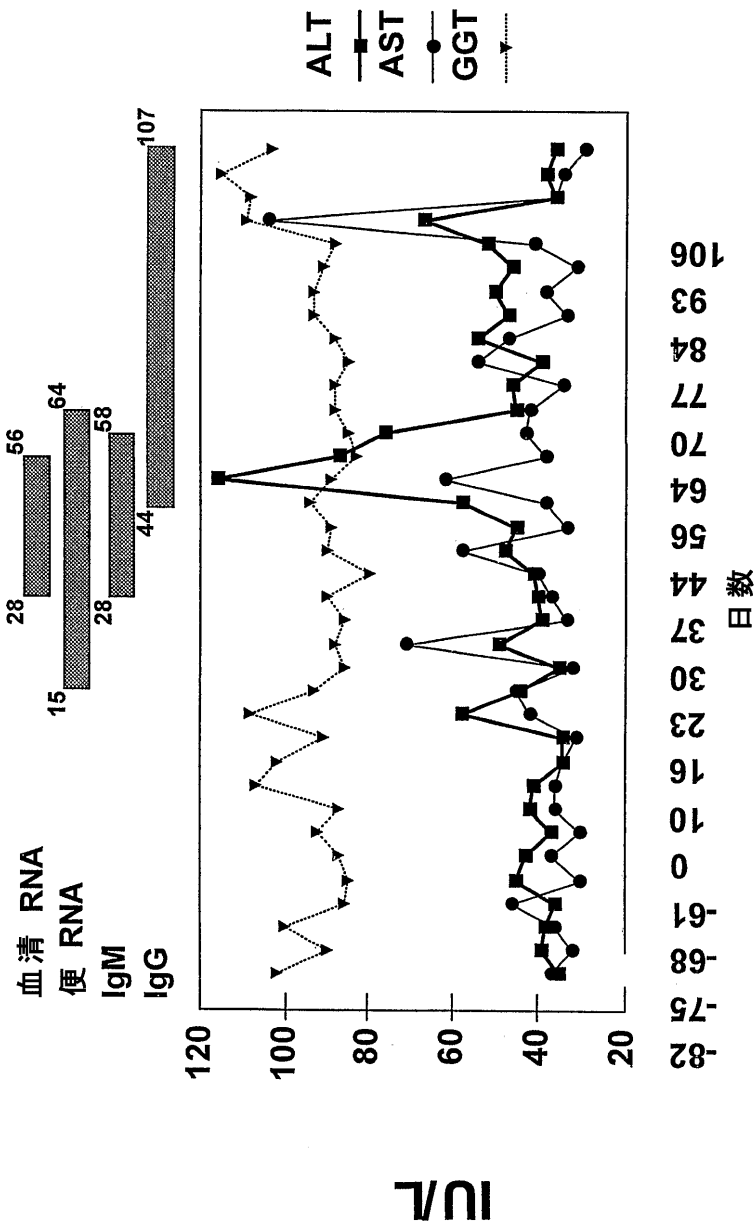


Figure 7

Z12

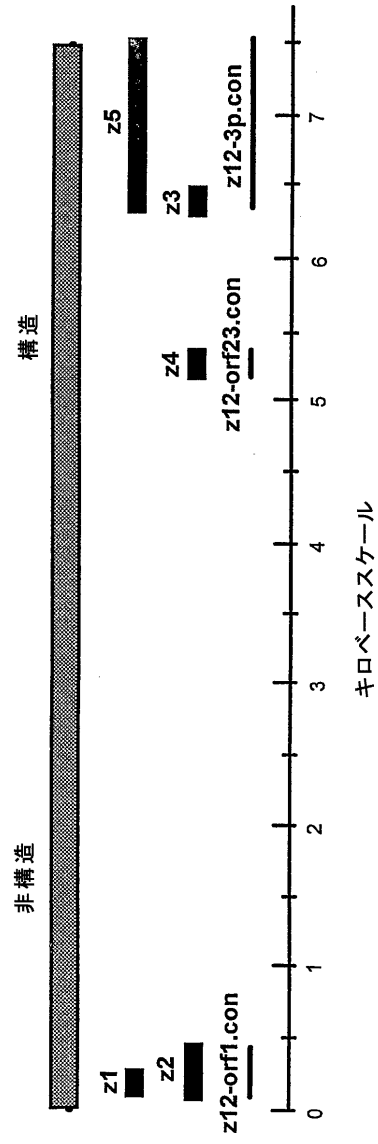


Figure 8

FIGURE 9B

5p.pile{hpesvp}	201	CCCCGAGGTT	TTCTGGAATC	ATCCCATCCA	GGGTGTCATC	CATAACGAGC	250
5p.pile{hpeuigh}		CCCCGAGGTT	TTCTGGAACC	ACCCCATCCA	GGGTGTCATC	CATAATGAGC	
5p.pile{hpea}		CCCCGAGGTT	TTCTGGAACC	ATCCCATCCA	GGGTGTTATC	CATAATGAGC	
5p.pile{840455p}		CCCTGAGGTA	CTTTGGAATC	ACCCATATCCA	GCGGGTTATA	CATAATGAAT	
5p.pile{hpenssp}		TCCTGAGGTT	TTTTGGAATC	ACCCGATTCA	ACGTGTTATA	CATAATGAGC	
コンセンサス		-CC-GAGGT-	-T-TGGAA-C	A-CC-AT-CA	-CG-GT-AT-	CATAA-GA--	
5p.pile{hpesvp}	251	TGGAGCTTTA	CTGCCGCGCC	CGCTCCGGCC	GCTGTCTTGA	AATTGGCGCC	300
5p.pile{hpeuigh}		TGGAGCTTTA	CTGTGCGGCC	CGCTCCGGCC	GCTGCCCTTGA	AATTGGTGCC	
5p.pile{hpea}		TGGAGCTTTA	CTGTGCGGCC	CGCTCCGGCC	GCTGCCCTCGA	AATTGGTGCC	
5p.pile{840455p}		TAGAACAGTA	CTGCCGGGCT	CGGGCTGGTC	GTTGCTTGA	GTTGGAGCT	
5p.pile{hpenssp}		TTGAGCAGTA	TTGCCGTGCT	CGCTCGGGTC	GCTGCCCTGA	GATTGGAGCC	
コンセンサス		T-GA-C--TA	-TG-CG-GC-	CG--C-GG-C	G-TG--T-GA	--TTGG-GC-	
5p.pile{hpesvp}	301	CATCCCCGCT	CAATAAATGA	TAATCCTAAT	GTGGTCCACC	GCTGCTTCCT	350
5p.pile{hpeuigh}		CACCCTCGCT	CAATAAACGA	CAATCCTAAT	GTGGTCCACC	GCTGCTTCCT	
5p.pile{hpea}		CACCCCGGCT	CAATAAATGA	CAATCCTAAT	GTGGTCCACC	GTTGCTTCCT	
5p.pile{840455p}		CACCCAAGAT	CCATTAATGA	CAACCCCAAC	GTTCTGCATC	GGTGTTCCT	
5p.pile{hpenssp}		CACCCACGCT	CCATTAATGA	TAATCCTAAT	GTCCTCCATC	GCTGCTTCCT	
コンセンサス		CA-CC--G-T	C-AT-AA-GA	-AA-CC-AA-	GT--T-CA-C	G-TG-TT-CT	

【図 9 C】

FIGURE 9C

5p.pile{hpesvp}	351	CCGCCCTGTT	GGCGTGATG	TTCAGCGCTG	GTATACTGCT	400	CCCACTCGCG
5p.pile{hpeuigh}		CCGCCCTGCC	GGCGTGATG	TTCAGCGCTG	GTATACTGCT		CCTACCCGCG
5p.pile{hpea}		CCGTCCCTGCC	GGCGTGATG	TTCAGCGCTG	GTATACTGCC		CCTACCCGCG
5p.pile{840455p}		TAGACCGGTT	GGCCGAGATG	TTCAGCGCTG	GTACTCTGCC		CCACCCGCG
5p.pile{hpenssp}		CCACCCCGTC	GGCCGGGATG	TTCAGCGCTG	GTACACAGCC		CCGACTAGGG
コンセンサス		-----CC-G--	GG-CG-GATG	TTCAGCGCTG	GTA--C-GC-		CC-AC--G-G
5p.pile{hpesvp}	401	GGCCGGCTGC	TAAATTGCCG	CGTTCCGCGC	TGCGCGGGCT	450	TCCCGCTGCT
5p.pile{hpeuigh}		GGCCGGCTGC	TAAATTGCCG	GGTTCCGCAC	TGCGCGGGCT		CCCCGCTGCT
5p.pile{hpea}		GGCCGGCTGC	TAAATTGCCG	CGTTCCGCGC	TGCGCGGGCT		CCCCGCTGCT
5p.pile{840455p}		GCCCTGcGGc	TAAATTGCCG	cGcTcGCGT	TGCGTGGTCT		CCCCCGCGCT
5p.pile{hpenssp}		GACCTGCGGC	GAACTGTcGC	CGCTCGGCAC	TTCGTGGTCT		GCCACcAGCC
コンセンサス		G-CC-GC-GC	-AA-TG-CG-	-G-TC-GC--	T-CG-GG-CT		-CC--C-GC-
5p.pile{hpesvp}	451	GACCGCACTT	ACTCCCTCGA	CGGGTTTCT	GGCTGTAACT	500	TTCCCGCCGA
5p.pile{hpeuigh}		GACCGCACTT	ACTGCTTCGA	CGGGTTTCT	GGCTGTAACT		TTCCCGCCGA
5p.pile{hpea}		GACCGCACTT	ACTGCTTCGA	CGGGTTTCT	GGCTGTAACT		TTCCCGCCGA
5p.pile{840455p}		GACCGCACTT	ACTGCTTTGA	TGGATTCTCC	CGTTGTGCTT		TTGCTGCAGA
5p.pile{hpenssp}		GACCGCACTT	ACTGTTTGA	TGGCTTTGCC	GGCTGCCGTT		TTGCCGCCGA
コンセンサス		GACCGCACTT	ACTG--T-GA--GG-	TT--C-	-G-TG-----T		TT-C-GC-GA

【図9D】

FIGURE 9D

3p.pile{hpea}	1451	ACTGAGTCAG	TGAAGCCAGT	GCTTGACCTG	ACAAATTCAA	TTCTGTGTCTG	1500
3p.pile{hpeuigh}		ACTGAGTCGG	TGAAGCCAGT	GCTCGACTTG	ACAAATTCAA	TCCTGTGTCTG	
3p.pile{hpesvp}		ACTGAGTCAG	TAAACCAGT	GCTCGACTTG	ACAAATTCAA	TCTGTGTCTG	
3p.pile{hpenssp}		ACAGAGTCTG	TAAAGCCTAT	ACTTGACCTT	ACACACTCAA	TTATGACACCG	
3p.pile{840453p}		ACAGAGACTA	TAAACCTGT	ACTTGATCTC	ACAAATTCCA	TCATACAGCG	
コンセンサス		AC-GAG-C--	T-AA-CC--T	-CT-GA--T-	ACA-A-TC-A	T--T	
3p.pile{hpea}	1501	GGTGGAATGA	ATAACATGTC	TTTTGCTGCG	CCCATGGGTT	CGCGACCATG	1550
3p.pile{hpeuigh}		GGTGGAATGA	ATAACATGTC	TTTTGCTGCG	CCCATGGGTT	GGCGACCATG	
3p.pile{hpesvp}		GGTGGAATGA	ATAACATGTC	TTTTGCTGCG	CCCATGGGTT	CGCGACCATG	
3p.pile{hpenssp}		GTCTGAATGA	ATAACATGTC	TTTTGCTGCG	CCCATGGGTT	CGCGACCATG	
3p.pile{840453p}		GGTGGAATGA	ATAACATGTC	TTTTGCTGCG	CCCATGGGTT	CGCGACCATG	
コンセンサス		GGTGGAATGA	ATAACATGTC	TTTTGCTGCG	CCCATGGGTT	CGCGACCATG	
3p.pile{hpea}	1551	CGCCCTCGGC	CTATTTTGCT	GTTGCTCCTC	ATGTTTCTGC	CTATGCTGCC	1600
3p.pile{hpeuigh}		CGCCCTCGGC	CTATTTTGCT	GTTGCTCCTC	ATGTTTCTGC	CTATGCTGCC	
3p.pile{hpesvp}		CGCCCTCGGC	CTATTTTGCT	GTTGCTCCTC	ATGTTTCTGC	CTATGCTGCC	
3p.pile{hpenssp}		CGCCCTAGGC	CTCTTTTGCT	GTTGTTCTCTC	TTGTTTCTGC	CTATGTTGCC	
3p.pile{840453p}		CGCCCTAGGC	CTGTTCTGTT	GTTGTTCTCTC	ATGTTTCTGC	CTATGCTGCC	
コンセンサス		CGCCCT-GG-	CT-TT-TG-T	G-TG-TCCTC	-TGTTT-TGC	CTAT--TGCC	

【図 9 E】

FIGURE 9E

3p.pile{hpea}	CGGCCACCG	CCCGGTCAGC	CGTCTGGCCG	CCGTCGTGGG	CGGCGCAGCG	1650
3p.pile{hpeuigh}	CGGCCACCG	CCCGGTCAGC	CGTCTGGCCG	CCGTCGTGGG	CGGCGCAGCG	
3p.pile{hpesvp}	CGGCCACCG	CCCGGTCAGC	CGTCTGGCCG	CCGTCGTGGG	CGGCGCAGCG	
3p.pile{hpenssp}	CGGCCACCG	ACCGGTCAGC	CGTCTGGCCG	CCGTCGTGGG	CGGCGCAGCG	
3p.pile{840453p}	CGGCCACCG	GCGGTCAGC	CGTCTGGCCG	TCGCCGTGGG	CGGCGCAGCG	
コンセンサス	CGGCCACCG	-CCGGTCAGC	CGTCTGGCCG	-CG-CGTGGG	CGGCGCAGCG	
3p.pile{hpea}	GCGGTTCCGG	CGGTGGTTTC	TGGGGTGACC	GGGTTGATTTC	TCAGCCCTTC	1700
3p.pile{hpeuigh}	GCGGTTCCGG	CGGTGGTTTC	TGGGGTGACC	GGGTTGATTTC	TCAGCCCTTC	
3p.pile{hpesvp}	GCGGTTCCGG	CGGTGGTTTC	TGGGGTGACC	GGGTTGATTTC	TCAGCCCTTC	
3p.pile{hpenssp}	GCGGTACCGG	CGGTGGTTTC	TGGGGTGACC	GGGTTGATTTC	TCAGCCCTTC	
3p.pile{840453p}	GCGGTGCCGG	CGGTGGTTTC	TGGAGTGACA	GGGTTGATTTC	TCAGCCCTTC	
コンセンサス	GCGGT-CCGG	CGGTGGTTTC	TGG-GTGAC-	GCGGTTGATTTC	TCAGCCCTTC	
3p.pile{hpea}	GCAATCCCCCT	ATATTCAATCC	AACCAACCCC	TTGCCCCCCC	ATGTCACCGC	1750
3p.pile{hpeuigh}	GCAATCCCCCT	ATATTCAATCC	AACCAACCCC	TTGCCCCCCC	ATGTCACCGC	
3p.pile{hpesvp}	GCAATCCCCCT	ATATTCAATCC	AACCAACCCC	TTGCCCCCCC	ATGTCACCGC	
3p.pile{hpenssp}	GCAATCCCCCT	ATATTCAATCC	AACCAACCCC	TTGCCCCCAG	ACGTTGCCGC	
3p.pile{840453p}	GCCCTCCCCCT	ATATTCAATCC	AACCAACCCC	TTGCCCCGCC	ATGTCGTTTC	
コンセンサス	GC-TCGCCCT	ATATTCAATCC	AACCAACCCC	TT-GCC-C-G	A-GT-----C	

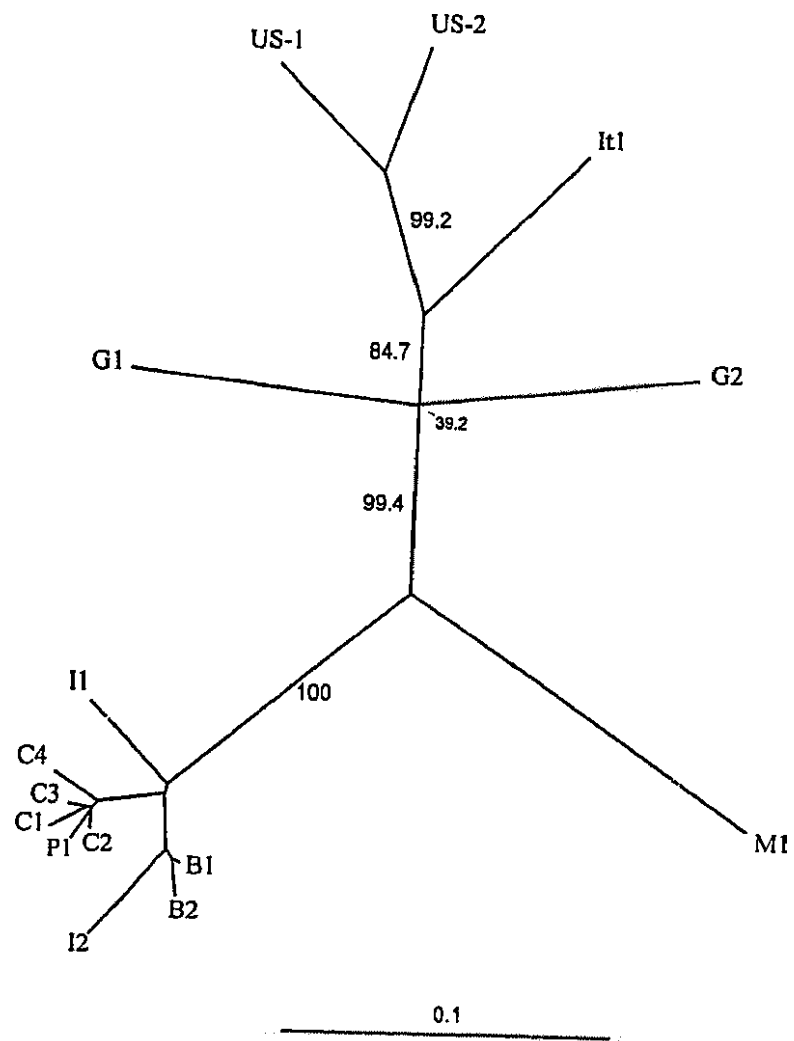
FIGURE 9F

3p.pile{hpea}	2651	AGCGCTTACC	CTGTTTAACC	TTGCTGACAC	CCTGCTTGGC	GGTCTACCGA	2700
3p.pile{hpeuigh}		AGCGCTTACC	CTGTTTAACC	TTGCTGACAC	CCTGCTTGGC	GGTCTACCGA	
3p.pile{hpesvp}		AGCCCTCACC	CTGTTCAACC	TTGCTGACAC	TCTGCTTGGC	GGCCTGCCGA	
3p.pile{hpenssp}		AGCTCTAACA	TACTTTAACC	TTGCTGACAC	GCTCCTCGGC	GGGCTCCCGA	
3p.pile{840453p}		TGCCCTGACT	CTGTTTAATC	TTGCTGATAC	GCTTCTTGGT	GGTTTACCGA	
コンセンサス		-GC-CT-AC-	-T--T-AA-C	TTGCTGA-AC	-CT-CT-GG-	GG--T-CC-	
3p.pile{hpea}	2701	CAGAAATTGAT	TTCGTCGGCT	GGTGGCCAGC	TGTTCTACTC	TCGCCCCGTC	2750
3p.pile{hpeuigh}		CAGAAATTGAT	TTCGTCGGCT	GGTGGCCAGC	TGTTCTACTC	TCGCCCCGTC	
3p.pile{hpesvp}		CAGAAATTGAT	TTCGTCGGCT	GGTGGCCAGC	TGTTCTACTC	CCGTCCCGTT	
3p.pile{hpenssp}		CAGAAATTAAT	TTCGTCGGCT	GGCGGGCAAC	TGTTTTATTTC	CCGCCCCGTT	
3p.pile{840453p}		CAGAAATTGAT	TTCGTCGGCT	GGGGGTCAAC	TGTTTTACTC	CCGCCCCGTT	
コンセンサス		CAGAAATTGAT	TTCGTCGGCT	GGGGGTCAAC	TGTT-TA-TC	-CG-CC-CT-	
3p.pile{hpea}	2751	GTCTCAGCCA	ATGGCGAGCC	GACTGTTAAG	CTGTATACAT	CTGTGGAGAA	2800
3p.pile{hpeuigh}		GTCTCAGCCA	ATGGCGAGCC	GACTGTTAAG	CTGTATACAT	CTGTAGAGAA	
3p.pile{hpesvp}		GTCTCAGCCA	ATGGCGAGCC	GACTGTTAAG	TTGTATACAT	CTGTAGAGAA	
3p.pile{hpenssp}		GTCTCAGCCA	ATGGCGAGCC	AACCGTGAAG	CTCTATACAT	CAGTGGAGAA	
3p.pile{840453p}		GTCCTCGCCA	ATGGCGAGCC	AACAGTAaAG	TTATACACAT	CTGTTGAGAA	
コンセンサス		GTCCTCGCCA	ATGGCGAGCC	-AC-GT-AAG	-T-TA-ACAT	C-GT-GAGAA	

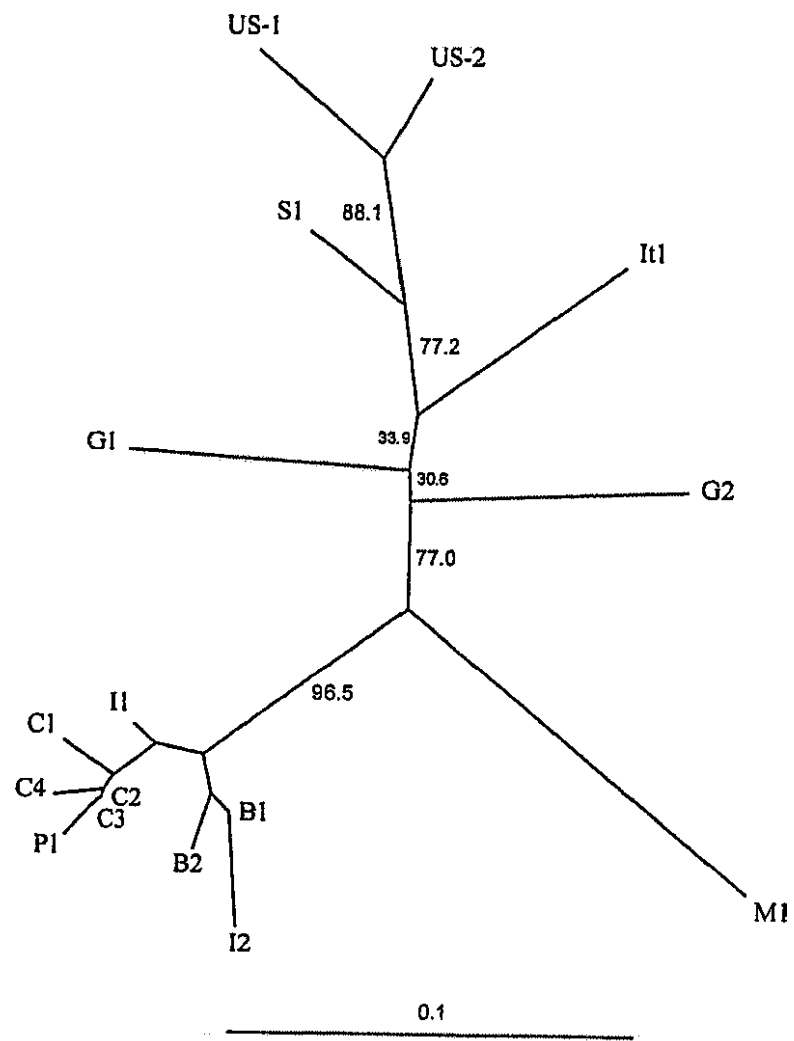
FIGURE 9G

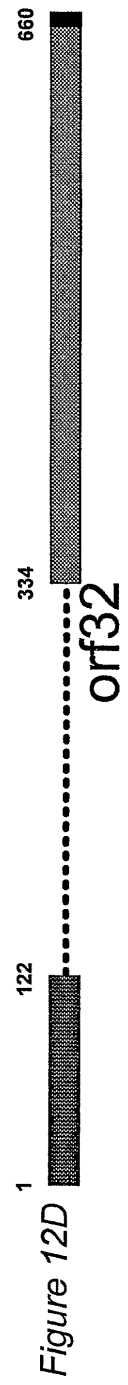
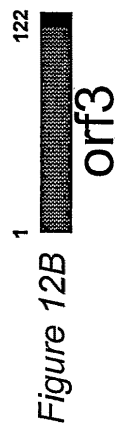
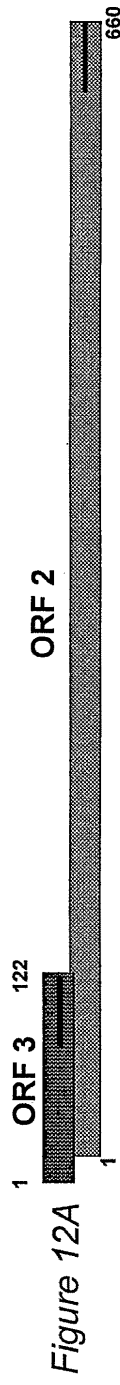
3p.pile{hpea} 3p.pile{hpeuigh} 3p.pile{hpesvp} 3p.pile{hpenssp} 3p.pile{840453p} コンセンサス	2801	TGCTCAGCAG	GATAAGGGTA	TTGCAATCCC	GCATGACATC	GACCTCGGGG	2850
		TGCTCAGCAG	GATAAGGGTA	TTGCAATCCC	GCATGACATC	GACCTCGGGG	
		TGCTCAGCAG	GATAAGGGTA	TTGCAATCCC	GCATGACATT	GACCTCGGAG	
		TGCTCAGCAG	GATAAGGGTG	TTGCTATCCC	CCACGATATC	GATCTTGGTG	
		TGCgCagCAA	gACAAGGGGca	TcacCaTTCC	ACACGACATA	gATTTAGGTG	
		TGC-CAGCA-	GA-AAGGG--	T--C-AT-CC	-CA-GA-AT-	GA--T-GG-G	
3p.pile{hpea} 3p.pile{hpeuigh} 3p.pile{hpesvp} 3p.pile{hpenssp} 3p.pile{840453p} コンセンサス	2851	AATCCCGTGT	AGTTATTTCAG	GATTATGACA	ACCAACATGA	GCAGGACCGA	2900
		AATCTCGAGT	TGTTATTTCAG	GATTATGACA	ACCAACATGA	GCAGGACCGG	
		AATCTCGTGT	GGTTATTTCAG	GATTATGATA	ACCAACATGA	ACAAGATCGG	
		ATTCCCGTGT	GGTCATTTCAG	GATTATGACA	ACCAGCATGA	GCAGGATCGG	
		ACTCCCGTGT	GGTTATCCAG	gattattgata	accagcacga	adaAgaTcgA	
		A-TC-CG-GT	-GT-AT-CAG	GATTATGCA-A	ACCA-CA-GA-	-CA-GA-CG-	

【図 1 0】

*Figure 10*

【図11】

*Figure 11*



【図13】

マカク 13906
HEV US-2 ORF 3 CKS - 29

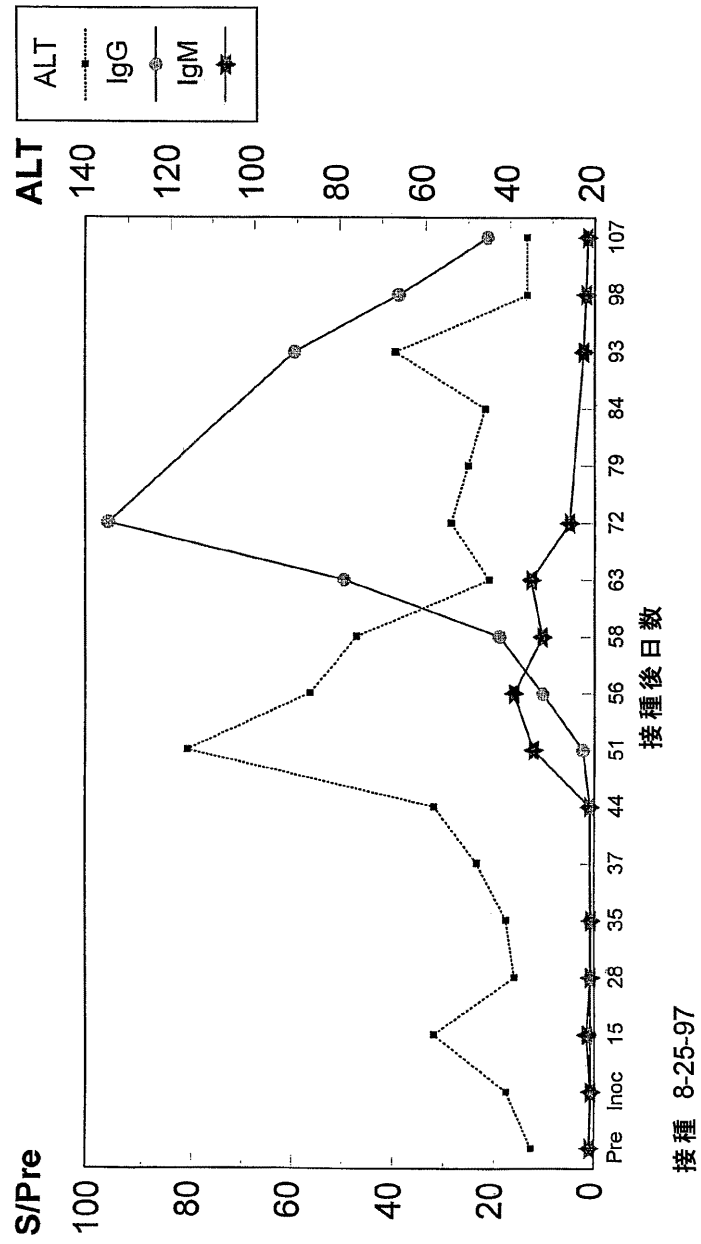


Figure 13

【図14】

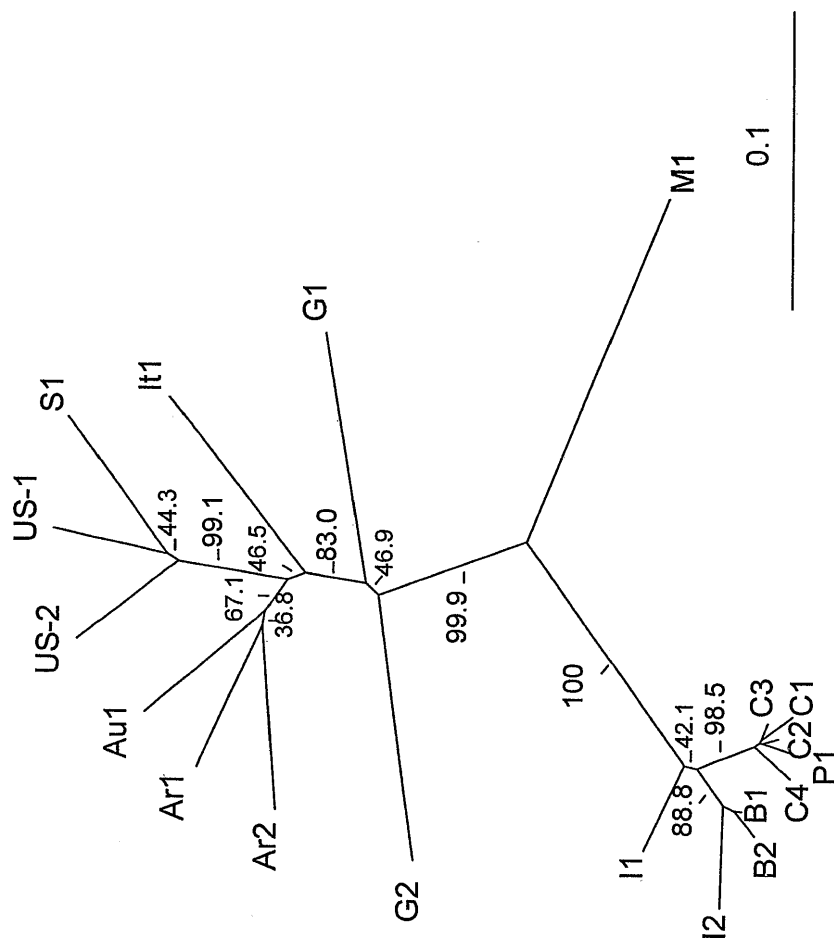


Figure 14

【図15】

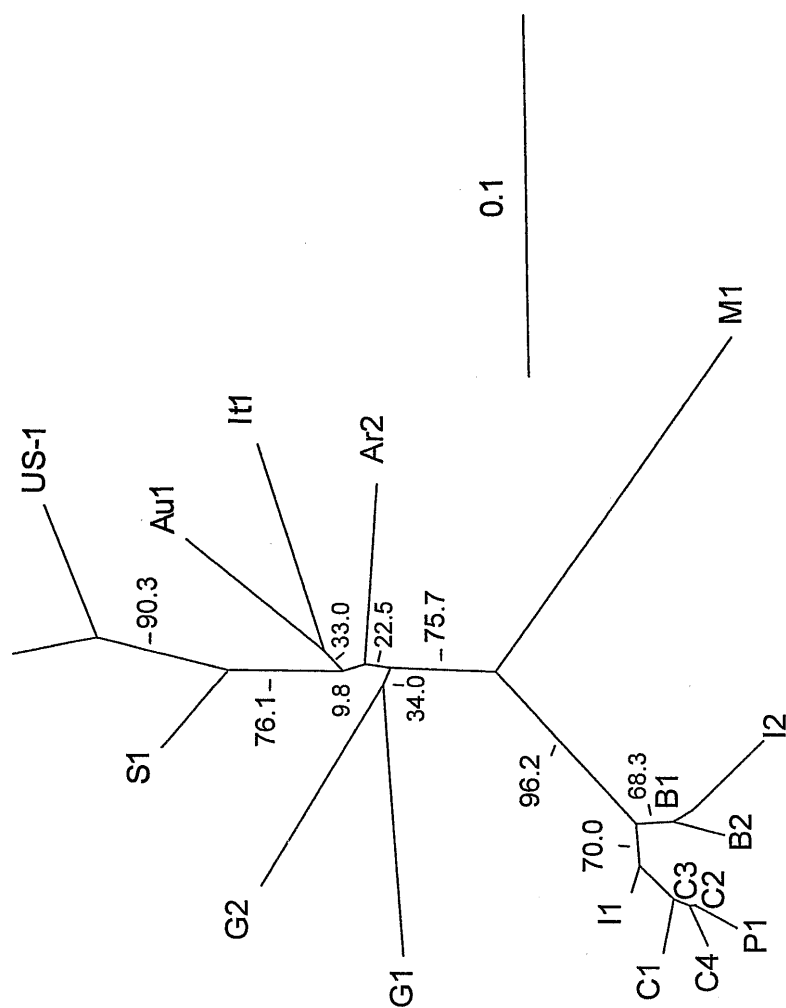


Figure 15

【図16】

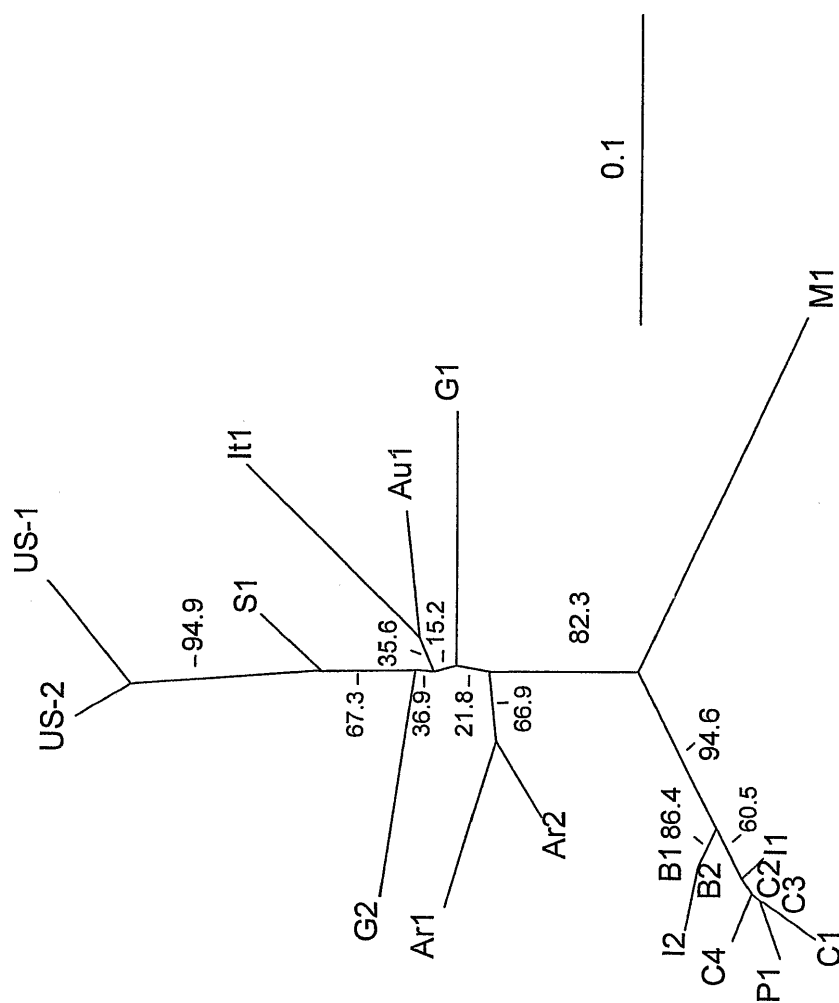


Figure 16

【国際調査報告】

INTERNATIONAL SEARCH REPORT

International Application No.
PCT/US 00/34420

A. CLASSIFICATION OF SUBJECT MATTER					
IPC 7	G01N33/576	G01N33/577	C12Q1/70	C07K14/08	C07K16/10
	C12N15/51	C12N5/10	A61K39/29	A61K39/42	A61P31/14
According to International Patent Classification (IPC) or to both national classification and IPC					
B. FIELDS SEARCHED					
Minimum documentation searched (classification system followed by classification symbols)					
IPC 7 G01N C07K					
Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched					
Electronic data base consulted during the international search (name of data base and, where practical, search terms used)					
EPO-Internal					
C. DOCUMENTS CONSIDERED TO BE RELEVANT					
Category *	Citation of document, with indication, where appropriate, of the relevant passages				Relevant to claim No.
X	WO 99 19732 A (ABBOTT LAB) 22 April 1999 (1999-04-22) the whole document				1-43
☐ Further documents are listed in the continuation of box C. ☒ Patent family members are listed in annex.					
* Special categories of cited documents : 'A' document defining the general state of the art which is not considered to be of particular relevance 'E' earlier document but published on or after the international filing date 'L' document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified) 'O' document referring to an oral disclosure, use, exhibition or other means 'P' document published prior to the international filing date but later than the priority date claimed 'T' later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention 'X' document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone 'Y' document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art 'Z' document member of the same patent family					
Date of the actual completion of the international search			Date of mailing of the international search report		
11 July 2001			18/07/2001		
Name and mailing address of the ISA European Patent Office, P.B. 5818 Patentlaan 2 NL - 2280 HV Rijswijk Tel. (+31-70) 340-2040, Tx. 31 651 epo nl Fax: (+31-70) 340-3010			Authorized officer: Gundlach, B		

INTERNATIONAL SEARCH REPORT

(information on patent family members)

International Application No

PCT/US 00/34420

Patent document cited in search report	Publication date	Patent family member(s)	Publication date
WO 9919732 A	22-04-1999	EP 1023605 A	02-08-2000

フロントページの続き

(51)Int.Cl. ⁷	識別記号	F I	テ-マコ-ト ⁷ (参考)
A 6 1 P 31/12		A 6 1 P 31/12	4 H 0 4 5
C 0 7 K 14/08		C 0 7 K 14/08	
	16/10		16/10
C 1 2 N 1/15		C 1 2 N 1/15	
	1/19		1/19
	1/21		1/21
	5/10	C 1 2 Q 1/68	A
	15/09	G 0 1 N 33/53	D
C 1 2 Q 1/68			33/576 Z
G 0 1 N 33/53	Z N A	C 1 2 N 15/00	Z N A A
		5/00	A
(72)発明者	アーカー , ジエイムズ・シー		
	アメリカ合衆国、イリノイ・60030、グレイズレイク、ノース・ホワイト・テイル・ドライブ・359		
(72)発明者	デサイ , スレツシュ・エム		
	アメリカ合衆国、イリノイ・60048、リバティビル、アミー・レイン・1408		
(72)発明者	ドーソン , ジョージ・ジエイ		
	アメリカ合衆国、イリノイ・60048、リバティビル、サウス・ダイヤモンド・ロード・914		
(72)発明者	マツシヤワー , アイサ・ケイ		
	アメリカ合衆国、イリノイ・60030、グレイズレイク、アーバー・ブルバード・18790		
F タ-ム(参考)	4B024 AA14 BA32 DA03 EA02 GA11		
	HA14		
	4B063 QA01 QQ10 QR32 QR55 QS34		
	4B065 AA93X AA96X AA96Y AA99Y		
	BA02 CA45		
	4C085 AA03 AA13 BA87 BB31 CC22		
	CC23 DD33 EE06 FF03		
	4C087 BB63 DA15 NA14 ZA75 ZB33		
	4H045 AA10 AA11 AA30 BA10 CA02		
	DA75 DA86 EA22 EA28 FA74		

专利名称(译)	<无法获取翻译>		
公开(公告)号	JP2003525428A5	公开(公告)日	2008-05-22
申请号	JP2001547552	申请日	2000-12-19
[标]申请(专利权)人(译)	雅培公司		
申请(专利权)人(译)	雅培制药		
当前申请(专利权)人(译)	雅培制药		
[标]发明人	シユロイダー・ジョージ・ジー アーカー・ジェームズ・シー デサイ・スレツシユ・エム ドーソン・ジョージ・ジェイ マツシャワー・アイサケイ		
发明人	シユロイダー, ジョージ・ジー アーカー, ジェームズ・シー デサイ, スレツシユ・エム ドーソン, ジョージ・ジェイ マツシャワー, アイサケイ		
IPC分类号	G01N33/569 A61K35/14 A61K39/00 A61K39/395 A61P31/12 C07K14/08 C07K16/10 C12N1/15 C12N1/19 C12N1/21 C12Q1/68 G01N33/53 G01N33/576 C12N15/09 C12N5/10		
CPC分类号	G01N33/5767 C12N2770/28122 C07K14/005 A61K39/00		
FI分类号	G01N33/569.L A61K35/14.Z A61K39/00.H A61K39/395.D A61K39/395.S A61P31/12 C07K14/08 C07K16/10 C12N1/15 C12N1/19 C12N1/21 C12Q1/68.A G01N33/53.D G01N33/576.Z C12N15/00.ZNA.A C12N5/00.A		
F-TERM分类号	4B024/AA14 4B024/BA32 4B024/DA03 4B024/EA02 4B024/GA11 4B024/HA14 4B063/QA01 4B063/QQ10 4B063/QR32 4B063/QR55 4B063/QS34 4B065/AA93X 4B065/AA96X 4B065/AA96Y 4B065/AA99Y 4B065/BA02 4B065/CA45 4C085/AA03 4C085/AA13 4C085/BA87 4C085/BB31 4C085/CC22 4C085/CC23 4C085/DD33 4C085/EE06 4C085/FF03 4C087/BB63 4C087/DA15 4C087/NA14 4C087/ZA75 4C087/ZB33 4H045/AA10 4H045/AA11 4H045/AA30 4H045/BA10 4H045/CA02 4H045/DA75 4H045/DA86 4H045/EA22 4H045/EA28 4H045/FA74		
优先权	09/468147 1999-12-21 US		
其他公开文献	JP2003525428A		

摘要(译)

公开了用于检测样品中US型或US亚型戊型肝炎病毒（包括其天然存在的变体）的存在的方法和组合物。特别地，本发明涉及对应于US型或US亚型戊型肝炎病毒的基因组的核酸序列，由所述病毒的基因组编码的氨基酸序列（包括表位序列），并特异性地与所述氨基酸序列结合。提供抗体。本发明进一步提供了使患者免疫所述病毒感染或治疗已经被所述病毒感染的患者的方法和组合物。

