Analysis of a strain of hepatitis E virus associated with acute liver failure

A thesis submitted in partial fulfilment for the award of doctor of philosophy

Dr Matthew Charles Donati

Centre for Hepatology

Royal Free and University College Medical School

University College London

Rowland Hill Street

London NW3 2PF

2001

ProQuest Number: U643457

All rights reserved

INFORMATION TO ALL USERS
The quality of this reproduction is dependent upon the quality of the copy submitted.

In the unlikely event that the author did not send a complete manuscript
and there are missing pages, these will be noted. Also, if material had to be removed,
a note will indicate the deletion.

ProQuest U643457

Published by ProQuest LLC(2016). Copyright of the Dissertation is held by the Author.

ProQuest LLC
789 East Eisenhower Parkway
P.O. Box 1346
Ann Arbor, MI 48106-1346

ABSTRACT

Hepatitis E virus (HEV) infection often results in a relatively mild, self-limiting hepatitis, although acute liver failure occurs in some patients, particularly women in the later stages of pregnancy. It is uncertain whether the outcome of infection is predominantly influenced by host or viral factors, or whether an interplay of both is important.

At the time of project inception, most HEV sequences in the nucleotide databases were from virus associated with uncomplicated cases of hepatitis E, and many were derived from virus passaged in monkeys. We set out to determine the nucleotide sequence of an HEV strain associated with acute liver failure in a male, using liver tissue as the source of viral RNA.

The consensus nucleotide sequence of this HEV strain was determined from standard length amplicon and full length genomic HEV PCR products. Comparison of the consensus sequence with published HEV sequences showed this strain of HEV to conform to the previously described genomic structure. The 5' non-coding region of this strain contained a unique nucleotide change at position 11. There were no obvious features that could account for the increased pathogenicity. The closest matches in the database were with genotype 1 HEV sequences.

Expression *in-vitro* of full length HEV genomes and individual ORF 2 and 3 generated proteins of the size expected from sequence data and published results. Transient expression in COS-7 cells failed to show the presence of HEV specific proteins by western blot or radioimmunoprecipitation or clear evidence of HEV RNA by northern blot. Northern blotting of total liver RNA did not indicate the presence of previously described subgenomic RNAs.

The outcome of infection with HEV is likely to be the result of the host immune response rather than altered viral pathogenicity, however further analysis of HEV strains associated with acute liver failure is required.

Acknowledgements

Dedicated to my family

INDEX

CHAPTER 1:   Introduction

CHAPTER 2:   Materials and Methods

CHAPTER 3: Results

CHAPTER 4:   Discussion

LIST OF FIGURES

9

LIST OF TABLES

APPENDICES

LIST OF ABBREVIATIONS

ALF	Acute liver failure
ALT	Alanine aminotransferase
AST	Aspartate transaminase
CPE	Cytopathic effect
Cyno	Cynomolgus macaque (*Macaca facicularis*)
ET-NANBH	Enterically transmitted non-A, non-B hepatitis
FHF	Fulminant hepatic failure
HAV	Hepatitis A virus
HBV	Hepatitis B virus
HCV	Hepatitis C virus
HDV	Hepatitis D virus (delta)
HEV	Hepatitis E virus
HPRNI	Human placental ribonuclease inhibitor
IEM	Immune electron microscopy
NANBH	Non-A, non-B hepatitis
NCR	Non-coding region
ORF	Open reading frame
PT-NANBH	Parenterally transmitted non-A, non-B hepatitis
RACE	Rapid amplification of cDNA ends
RDRP	RNA dependant RNA polymerase
VLP	Virus-like particle

CHAPTER ONE

INTRODUCTION

1.1 Overview of hepatitis viruses

The last twenty-five years have seen a great increase in the knowledge of viral hepatitis, a disease which affects hundreds of millions of people worldwide. Viral hepatitis causes significant morbidity and mortality, and a considerable financial burden, through both acute and chronic infections. The impact of viral hepatitis on particular countries depends upon the relative prevalence of the different agents and their ability to cause clinically apparent disease or long term sequelae.

Prior to the advent of serological and molecular assays, the diagnosis of hepatitis was based upon clinical and epidemiological features. The development of specific tests for markers of hepatitis A and B infection allowed screening of patients, the results of which indicated that at least two additional forms of hepatitis existed; these were descriptively termed enterically transmitted and parenterally transmitted non-A, non-B hepatitis. The main causative agents for these are now known to be the hepatitis E virus (Reyes *et al.*, 1990) and the hepatitis C virus (Choo *et al.*, 1989) respectively. Currently there are five characterised, unrelated, hepatitis viruses, namely hepatitis A, B, C, delta and E viruses. Many other viruses can cause hepatitis, for example Epstein-Barr virus and cytomegalovirus, however with these agents hepatitis is not the commonest feature of infection.

1.1.1 Hepatitis A virus

The causative agent of what was formerly called infectious hepatitis is now known to be the hepatitis A virus, a hepatovirus within the *Picornaviridae* (Melnick, 1992). The particle is between 25-28 nm in diameter, with no envelope. It has a positive single-stranded, polyadenylated RNA genome of around 7.5 kb, including a 5' non-coding region of 733 nucleotides and a 3' non-coding region of around 60 nucleotides (Cohen

et al., 1987). The single ORF encodes a polyprotein which is proteolytically processed. There are three major structural polypeptides, with a fourth, short, polypeptide predicted by sequence data. Hepatitis A virus can be grown in cell culture (Provost *et al.*, 1979), with repeated passage leading to adaptation and sometimes attenuation, linked to nucleotide changes in the 5' NCR and non-structural regions of the genome.

Infection with hepatitis A virus is predominantly by the faecal-oral route leading to sporadic cases and epidemics. In many resource poor countries the majority of the population are infected in childhood. In more developed regions, the improved sanitation has led to a shift in herd immunity and a trend towards more adult infections, with only between 5 to 20% of young adults seropositive for anti-HAV IgG. This shift in age at exposure is important since infection in childhood is usually asymptomatic, whereas adults often develop jaundice, with up to 1% suffering acute liver failure. Although rare, both protracted and relapsing hepatitis A have been described (Schiff, 1992). There is only one serotype of HAV and infection appears to confer lifelong immunity. Hepatitis A may be suspected on the basis of clinical details but diagnosis usually involves testing for specific IgM antibody.

Initially, specific prophylaxis against hepatitis A required the use of normal human immunoglobulin. The titre of HAV antibody in European donor derived immunoglobulin has declined in parallel with the incidence of the disease, possibly affecting its efficacy. A major advance in the control of hepatitis A has been the production of an effective vaccine based on killed whole virus (Andre *et al.*, 1990).

1.1.2 Hepatitis B virus

The path to discovery of the hepatitis B virus began in 1963 with the detection of an antigen in blood (Australia antigen, now known to be hepatitis B surface antigen), later shown to be associated with serum hepatitis (Blumberg *et al.*, 1967). Visualisation of HBV particles (Dane *et al.*, 1970) and cloning of the genome (Galibert *et al.*, 1979) followed. HBV is a member of the Hepadnaviridae, the 42 nm virion comprising a glycoprotein envelope and a nucleocapsid containing a DNA polymerase and a partially double stranded DNA genome, as well as the core and e antigens. The genome is around 3182 bp and codes for four overlapping genes, namely surface, core, polymerase and X. The entire surface ORF includes pre-S1, pre-

S2 and S domains, each with a separate initiation codon; the major surface protein is translated from the S domain. The core protein is translated from the second of two initiation codons within the core ORF. The e antigen is derived from proteolytic processing of the translation product of the first core ORF initiation codon. The polymerase ORF encodes the viral polymerase, which possesses both DNA and RNA dependent activities. The X ORF encodes a transcriptional transactivator.

Subtypes based upon a type common antigen (*a*) and two pairs of mutually exclusive determinants (*d/y and w/r*) on the surface antigen are associated with geographical regions. Subtype *adw* predominates in northern Europe. The different major subtype groupings were further divided, to give nine HBsAg subtypes. Classification based upon nucleotide similarity has been described (Lindh *et al.*, 1998). There are now seven HBV genotypes, from A to G (Norder *et al.,* 1994) and genotype nomenclature is now routinely used.

Hepatitis B occurs worldwide, with 350 million individuals with chronic infection and 2 billion people with serological evidence of past or current infection. Around 1 million people die each year as a result of acute hepatitis B or the effects of chronic infection. Regions may be divided according to prevalence; the UK has a low prevalence with a carrier rate of <0.1% population. Transmission is parenteral, with infection in areas of high prevalence, such as South East Asia and Sub-Saharan Africa, resulting predominantly from mother to infant transmission (vertical) and child to child (horizontal) transmission.

Infection is often subclinical, or it may manifest as a non-specific illness without jaundice or as a typical hepatitis, which can lead to acute liver failure. In addition, hepatitis B can be self-limited or become chronic. The outcome of infection is determined largely by age at exposure and immune status. Acute hepatitis is rare in neonates and infants, in whom infection often becomes chronic (70-90%) and a high proportion of these carriers (30-50%) suffer chronic liver disease and an increased risk of cirrhosis and hepatocellular carcinoma. Adult infection is more often symptomatic (30-40%), with less than 5% of immunocompetent individuals becoming chronic carriers and less than 1% developing fulminant hepatitis. Those individuals who are positive for e antigen or have high levels of HBV DNA in their plasma are at highest risk of liver disease. Immune deficiency predisposes to chronic carriage.

Treatment of carriers has been partially successful using both immune modulation and antiviral agents. Most experience has been with α interferon and the nucleoside

analogue, lamivudine (3TC). However, only a proportion of patients will lose e antigen or show a sustained reduction in HBV DNA levels (Lau *et al.*, 1997). The development of clinically significant resistance mutations to lamivudine is common if therapy is prolonged or repeated (Lau *et al.*, 2000). These mutations have been shown to reduce replication competence of HBV using *in-vitro* models (Ling and Harrison, 1999). Liver transplantation may be required for those with end-stage or fulminant disease, but graft re-infection may occur despite the use of antiviral agents and hepatitis B immunoglobulin prophylaxis.

Passive immunisation with specific hepatitis B immunoglobulin is available for use in single exposure events. A safe and effective HBV vaccine comprising plasma derived surface antigen has been developed (Buynak *et al.*, 1976), and this has been available since 1982. This was followed in 1985 by the introduction of vaccine based upon part of the surface antigen expressed in yeast (McAleer *et al.*,1984).

Several mutations within the HBV pre-core (Carman *et al.*, 1989) and surface genes (Harrison *et al.*, 1991) have been characterised. These probably arise as a result of selection pressure from host immunity or therapeutic measures. Pre-core mutations leading to non-production of e antigen may be associated with more severe disease (Inoue *et al.*, 1998). Since the e antigen is used as an indicator of infectivity, such variants may have implications for assessing transmission risk in health-care settings. Antigenic changes in the *a* determinant of the surface gene may allow infection despite the presence of vaccine induced antibody (Harrison *et al.*, 1991). The impact of these mutations on various aspects of hepatitis B infection has yet to be fully determined.

1.1.3 Hepatitis D (delta) virus

Delta hepatitis was first suggested after the detection of a novel antigen in hepatocytes of patients with hepatitis B (Rizetto *et al.*, 1977). HDV is a satellite virus (Gerin, 1994) which uses HBV as the helper agent, therefore it is only found in patients already infected, or co-infected, with hepatitis B. HDV is small, with a 1.7 kb RNA genome and two structural proteins (delta antigen) surrounded by an envelope of hepatitis B surface antigen. The hepatitis B surface antigen envelope protects the HDV RNA-delta antigen complex, and is required for entry into, and egress from, the

hepatocytes. HDV can replicate in a cell without the presence of HBV, but infectious HDV particles cannot be formed in its absence.

Large variations in geographical prevalence exist for HDV, which infects around 15 million individuals worldwide, many of them suffering severe liver disease. The clinical spectrum of infection is wide, including subclinical and fulminant courses and an accelerated progression to cirrhosis (reviewed by Hadziyannis, 1997). Co-infections with HBV and super-infections lead to a higher frequency of severe liver disease than that seen with HBV alone. Diagnostic tests for HDV infection include antibody, antigen and RNA (hybridisation and PCR) assays. Therapy with interferon can be effective in a limited proportion of patients but liver transplantation may be necessary in some. Immunisation against hepatitis B confers protection against HDV.

1.1.4 Hepatitis C virus

Hepatitis C infection is a worldwide major health concern, with an estimated 100 million people affected. The original term encompassing HCV, parenterally transmitted non-A, non-B hepatitis virus, reflected epidemiological observations and serological exclusion of other agents. There is a seroprevalence of around 1-2% in many developed countries, with higher rates in eastern Europe and Africa (Alter, 1995); Egypt has the highest rate at around 15-20%, possibly due to iatrogenic transmission during mass antischistosomal therapy (Frank *et al.*, 2000). Some modes of transmission are well recognised, such as transfusion, transplantation, injecting drug use, and needlestick injury, but many infections have no clear risk associations. Characterisation of HCV began with isolation of a cDNA clone in 1989 (Choo *et al.*, 1989) and it is now classified as a member of the *Flaviviridae*. The virus particle is enveloped with a positive-sense, single-stranded genome of around 9.4 kb. There is no polyadenylation at the 3' end, however, there are short 5' and 3' non-coding regions at the genome termini. There is a single ORF and polypeptides are produced by post-translational cleavage. Sequence variability has led to the description of different genotypes of HCV (Simmonds *et al.*, 1994), of which there are at least 11, and these are further subtyped on the basis of genetic relatedness (Davidson *et al.*, 1995). These exhibit geographical distribution, with 1a and 1b accounting for most infections within Europe.

Acute HCV infection rarely results in specific symptoms or jaundice. Infection becomes chronic in the majority (around 85%) and often first manifests when the complications of chronic liver disease arise. Nearly all patients with persistent HCV infection develop histological evidence of chronic hepatitis. Progression to cirrhosis may occur in up to 20% of patients within 10-20 years, and this is linked to a higher risk of hepatocellular carcinoma.

Diagnosis of HCV infection initially relies upon detection of serum antibody. Commonly used EIAs employ a mixture of antigens representing structural and non-structural regions of the virus. HCV antibody assays now routinely in use are in their third generation, despite this, there is still a relatively high false positive rate when screening populations at low risk for HCV infection. Supplementary testing by immunoblot is required to confirm reactivity in the EIAs. The time between infection and antibody detection may be 2-3 months in immunocompetent patients, and antibody may not be detectable at all in the immunosuppressed. In this situation, the detection of viraemia (e.g. antigen detection as part of a fourth generation assay, or by RT-PCR) is appropriate.

The management of chronic hepatitis C requires assessment of the degree of liver disease, followed by antiviral therapy, if appropriate (Dusheiko et al., 1996). Even after the most efficacious regimen comprising interferon alpha and ribavirin (Brillanti et al., 1994, Reichard et al., 1998), recurrence of viraemia and liver disease is common. The introduction of pegylated interferon with improved pharmacokinetics, may lead to better therapy response rates (Lindsay et al., 2001). The predictors of poor treatment outcome include a high HCV RNA level (>2 million copies per ml) and infection with genotype 1. Those patients with severe liver impairment may require organ transplantation, and although early survival is good, almost all grafts are re-infected and the long term effects of this are not fully quantified. Currently, there is no effective vaccine for HCV and specific immunoglobulin is not protective.

1.1.5 Hepatitis E virus

The development of serological tests for hepatitis A and B infection led to the description of enterically transmitted non-A, non-B hepatitis (ET-NANBH), the existence of which was already suspected on epidemiological grounds. Outbreaks and sporadic cases of ET-NANBH were described principally from developing or resource poor countries, with cases from developed countries generally resulting from travel to such areas. Several key features of ET-NANBH became apparent, notably the high mortality when infection occurred in the last trimester of pregnancy, and the low secondary household transmission rate. Molecular cloning of one of the agents associated with ET-NANBH, the hepatitis E virus (HEV), led to characterisation of the virus and the development of specific diagnostic tests. HEV is now known to be of wide geographic distribution and the major cause of sporadic and epidemic ET-NANBH in developing countries. The discovery, epidemiology, virology, and clinical features of HEV are discussed in detail later.

1.2 Other viral hepatitis agents

Serological and molecular based evidence indicates that many cases of viral hepatitis, including fulminant, acute, chronic, post-transfusion, and community acquired types, are not related to infection by hepatitis viruses A-E. Despite the ability to detect the known agents, at least 10-15% of acute clinical hepatitis in developing countries has an unknown aetiology. The proportion of cases of fulminant hepatitis in such areas with an unknown infective cause, is probably even higher. Around 20% of community acquired, sporadic hepatitis in Western Europe and the USA, may be due to uncharacterised infectious agents. Research directed toward discovery of the agents responsible for non A-E viral hepatitis has led to description of several candidate viruses, namely GBV-C, a paramyxovirus-like agent, a togavirus-like agent, TTV, TTV-like minivirus (TLMV), SANBAN virus, SEN virus, and sentinel virus (SNTV) (reviewed in Mushahwar, 2000). An adenovirus-like agent was proposed in 1994 as a cause of some cases of sporadic non-A, non-B viral hepatitis (Deka *et al.*, 1994). Subsequent reports have questioned the pathogenic role of this putative hepatitis F virus (Pillot *et al.*, 1997).

1.2.1 GB viruses

In 1967 serum from a surgeon (GB) with hepatitis was inoculated into a Tamarin (*Sanguinus sp.*) which then developed hepatitis (Deinhardt *et al.,*1967). Two new viruses were identified from tissue infected by passage in monkeys (Simons *et al.*, 1995). These viruses were called GBV-A and GBV-B. Assays to detect antibody against GBV-A and B identified immunoreactive samples amongst groups of normal blood donors, intravenous drug users, and individuals in West Africa. However, GBV-A or B genome sequences could not be detected in blood samples taken from these subjects. A third related virus was found in human sera by PCR screening using degenerate GBV sequence based primers, this was named GBV-C (Simons *et al.*, 1995b). Around the same time, HGV was characterised (Linnen *et al.*, 1996) and this is now known to be an independent isolate of GBV-C. GBV-A is now thought to be a naturally occurring Tamarin virus (Schlauder *et al.*, 1995); the origin of GBV-B is uncertain.

GBV-C is typical of the *Flaviviridae* with a positive sense RNA genome (around 9.5 kb) encoding a single polyprotein with the structural proteins at the amino terminus, however, the core gene has not yet been described. The role of GBV-C as a human pathogen is unclear (Alter *et al.*, 1997) and infection is often complicated by the presence of other hepatitis viruses. Most infections are asymptomatic and transient. GBV-C seems to infect cells of lymphoid origin, and detection of viral RNA in liver tissue may be from these (Tucker *et al.,* 2000). However, other studies have indicated that GBV-C is hepatotropic (Seipp *et al.*, 1999).

1.2.2 Paramyxovirus

A paramyxovirus-like agent has been suggested as a cause of giant-cell hepatitis (Phillips *et al.*, 1991) and possibly cirrhosis (Fimmel *et al.*, 1998). The possible viral nature of this form of hepatitis was enhanced by the observation that therapy with ribavirin may be beneficial (Durand *et al.*, 1997).

1.2.3 Togavirus

A further candidate hepatitis virus has been described in patients with non-A-E viral hepatitis and acute liver failure (Fagan *et al.*, 1992). Several patients had Togavirus-like particles seen in the native liver and in the grafted liver following recurrence of the illness after transplantation. This agent has not yet been characterised.

1.2.4 TTV

An unenveloped single stranded DNA virus, named TTV, has been proposed as a hepatitis agent (Nishizawa *et al.*, 1997). Analysis of the TTV genome has led to the suggestion that it is similar to the *Circoviridae*, and possibly a member of a new family, the *Circinoviridae* (Mushahwar *et al.*, 1999). The virus exhibits more than 30% diversity resulting in many genotypes, possibly as the result of homologous recombination. Transmission of this agent may be both enteral and parenteral (Okamoto *et al.*, 1998). The association of TTV with hepatitis is yet to be established, with conflicting results coming from different parts of the world (Tanaka *et al.*, 1998, Charlton *et al.*, 1998, Naoumov *et al.*, 1998, Simmonds *et al.*, 1998, Ikeda *et al.*, 1999), possibly resulting from the high prevalence of TTV in certain populations, for example up to 92% in Japan (Takahashi *et al.*, 1998). In addition, transmission to chimpanzees did not result in hepatitis (Mushahwar *et al.*, 1999) and TTV coinfection of HBV or HCV in humans did not affect the clinicopathological outcome of these infections (Kao *et al.*, 2000).

Attempts to detect TTV by PCR have led to the discovery of other, possibly related, viruses such as SANBAN (Hijikata *et al.*, 1998) and TLMV (Takahashi *et al.*, 2000). It is uncertain whether these viruses cause human disease.

1.2.5 Evidence for a non-A, non-E waterborne hepatitis agent

Although HEV is the major aetiological agent in many outbreaks of waterborne non-A hepatitis, it cannot be implicated in all the epidemics. In a retrospective seroepidemiological study (Arankalle *et al.*, 1994), 16 of 17 such outbreaks could be partially attributed to HEV. One epidemic, in the Andaman Islands off the East coast of India in 1987, was not serologically related to HEV when contemporary antibody

assays were used, and no other infectious or toxic cause for the hepatitis could be identified. The Andaman epidemic involved 307 cases of icterus. Of 81 patients from whom sera were available, 6 had acute HAV and the remainder had no obvious aetiology. Two features of the epidemic were unlike the typical HEV pattern- the attack rate was 31% and almost half of the cases were in patients under 15 years of age. This finding may indicate the existence of a new unidentified enterically transmitted agent or a variant HEV not detected by the assays available at the time of the study. These assays had been validated against a number of geographically distinct HEV outbreaks, however they did not contain an ORF 3 related antigen (see diagnosis section 1.3.9). A previous study had shown 3 of 5 samples from the Andaman epidemic to be HEV IgM positive by Western blot (Arankalle *et al.*, 1993) and therefore all samples were retested at lower dilution for IgG and IgM. From the results obtained the authors concluded that the initial reactivity in the earlier assay was probably non-specific. Applying more recent assays of varying format may help to clarify the situation.

1.3 Hepatitis E

Hepatitis E virus infection is an important public health concern in Asia, Africa and Central and South America. It often causes a relatively mild, self-limiting hepatitis, but acute liver failure and death occur in some patients. Treatment for liver failure requires intensive medical support and occasionally liver transplantation; these facilities are rarely readily available in areas where HEV is endemic. A clearer understanding of the factors that precipitate such a serious outcome of infection could lead to a reduction in its incidence.

1.3.1 Discovery of hepatitis E

Epidemics of hepatitis, resembling hepatitis A, and associated with contaminated drinking water, were reported in India and elsewhere from 1957 (Viswanathan, 1957; Sreenivasan, 1978; Khuroo, 1980). Clinical cases occurred in young adults in a setting where almost all adults had immunity to hepatitis A virus following childhood infection. In addition, the low number of secondary cases and the high mortality amongst women who were pregnant was atypical for outbreaks of hepatitis A. With

the advent of serological assays for HAV and HBV infection, these viruses were subsequently eliminated as aetiological agents (Wong *et al.*, 1980). Hence, this disease was originally named enterically transmitted non-A, non-B hepatitis (ET-NANBH).

In 1983 Prof M Balayan (at the Russian Academy of Medical Sciences, Moscow), who had serological evidence of immunity to hepatitis A, ingested clarified ET-NANBH associated stool and subsequently developed hepatitis. Faecally shed virus transmitted the disease to cynomolgus macaques (cynos) and reacted with sera from various victims of ET-NANBH (Balayan *et al.*, 1983). Virus like particles (VLPs) were visualised by immune electron microscopy (IEM) in Prof Balayan's acute phase stool and their physicochemical nature was found to be similar to that of the *Caliciviridae*. This represented the first description of transmission of the putative ET-NANBH agent to human and non-human primates. VLPs of a similar size to those observed by Balayan were reported by Kane in 1984 (Kane *et al.*, 1984) when convalescent serum was used in IEM of stool associated with an epidemic of ET-NANB in Nepal. Inoculation of monkeys (four marmosets and two chimpanzees) with a suspension of this stool led to an increase in liver enzyme activity in three of the four marmosets. Excretion of VLPs and production of antibody against the VLPs in the original inoculum occurred in one marmoset only. In 1987, Bradley and co-workers (Bradley *et al.*, 1987) reported serial transmission of VLPs associated with ET-NANBH to tamarins and cynomolgus macaques. The initial inoculum was derived from stools taken from epidemiologically and serologically defined cases of ET-NANBH from Pakistan and Burma. Biochemical and histological evidence of acute viral hepatitis was observed in the animals and serial passage of the disease in the macaques led to adaptation, as shown by a reduced incubation period and more severe hepatitis. These VLPs described in the early transmission studies were provisionally termed HEV.

Cloning and nucleotide sequencing of a section of HEV genome followed in 1990 (Reyes *et al.*, 1990). This involved extracting nucleic acid from bile obtained from a (third passage) cynomolgus macaque; the original source of virus was stool from a case of ET-NANB hepatitis in Burma. Bile provided high titre virus in a medium with a relatively simple nucleic acid content. cDNA prepared from both infected and uninfected bile was cloned into lambda gt10 bacteriophage and used in plus/minus screening. The first viral clone identified contained a motif representative of the RNA

dependent RNA polymerases found in all positive strand RNA viruses (the GDD putative active site). This clone was also found to hybridise to sequences of cDNA prepared from outbreak stool from different geographical regions. cDNA was amplified from these stools using a novel technique termed sequence independent single primer amplification (SISPA, Reyes *et al.*, 1992). An alternative screening method was used in addition to the above. cDNA from infected bile was also cloned into lambda gt11 and the expressed fusion proteins were screened using convalescent sera from an ET-NANB outbreak in Mexico. This approach identified two clones later found to localise within the 3' putative structural region of the genome. The first full length HEV nucleotide sequence was published in 1991 (Tam *et al.*, 1991) using primer extension to obtain the 5' end and detection of a polyadenylated clone indicating the 3' end. This was the Burmese prototype. There are now many published complete nucleotide sequences of human HEVs covering several geographic regions (see results section 3.6).

1.3.2 Hepatitis E virus biophysical characteristics

Hepatitis E virions have icosahedral symmetry, are usually between 32 and 34 nm in diameter by IEM, ranging from 27-38 nm, they are spherical, non-enveloped with spikes and indentations on the surface. The buoyant density is 1.29 g/ml. Those particles in the lower size range of 27-30 nm may be variants, or result from degradation of virions by proteolytic enzyme activity in the gut, or may be an artefact resulting from laboratory processing of stool. The virus is labile in high salt and degrades following freeze thawing cycles or prolonged storage unfrozen. These biophysical parameters initially resulted in the grouping of HEV within the family *Caliciviridae* (other members include canine and feline calicivirus, rabbit haemorrhagic disease virus, vesicular exanthem of swine virus, and human calicivirus eg Norwalk agent). HEV is currently unassigned by the International Committee on Taxonomy of Viruses (Green *et al.*, 2000), predominantly due to the lack of phylogenetic relatedness with other caliciviruses, the absence of a 5' VPg, and differences in replicative enzymes (Berke and Matson, 2000). This classification may change as new data are generated.

1.3.3 Hepatitis E virus genome and genotypes

Complete nucleotide sequences from several human isolates across various regions (including Burma, Mexico, Pakistan, China, India, USA) are now published, as well as a specific swine HEV. A comparison of these is given in the results section (chapter 3). The genetic divergence between strains generally reflects the geographical distance between the countries of origin. However, with the discovery of more diverse isolates, the original division into Old World and New World isolates is no longer valid. Despite the observed nucleotide sequence variation among isolates, cross-neutralisation studies suggest there is only one strain (Meng *et al.*, 1998b).

The Burmese strain prototype is described below, however differences in the length of some regions exist between this and other prototype strains and there is up to 8% nucleotide variation between Asian isolates and up to 25% between Asian, Mexican, North American, and European isolates.

The HEV genome (figure 1.1) differs from the *Caliciviridae* in its organisation (Green *et al.*, 2000) and by the suggested presence of a 5' cap instead of a genome linked polypeptide (VPg) (Kabrane-Lazizi *et al.*, 1999a; Magden *et al.*, 2001; Zhang *et al.*, 2001). The positive sense RNA genome is 7,194 nucleotides in length plus a poly A tail of approximately 300 nucleotides. ORF 1 contains the non-structural elements and is preceded by a 27 base non-coding stretch, before extending 5,079 bases prior to the stop codon. This 5' non-coding stretch may vary in length or composition between strains (see results, figure 3.6). ORF 2 begins 37 bases downstream of ORF 1 in the plus 1 frame and is 1,980 bases in length, followed by an apparently non-coding 68 base section and the poly A tail. The 3' non-coding section also varies in length across different HEV strains (see results, figure 3.7). ORF 3 is 369 bases and is in the plus 2 frame with respect to ORF 1. The 5' end of ORF 3 overlaps ORF 1 by one base and the 3' end overlaps ORF 2 by 328 bases.

Computer analysis of the nucleotide sequence of HEV ORF 1 has revealed sequences characteristic of (listed from 5' to 3') a methyltransferase, papain-like protease, RNA helicase and RNA dependent RNA polymerase (Koonin *et al.*, 1992). Analysis of several isolates has identified a hypervariable region around nucleotides 2100 to 2300 in ORF 1, possibly resulting from strong secondary structure in that area. This nucleotide analysis of ORF1 led to the suggestion that HEV is a member of an

alphavirus-like supergroup, with a close relationship to rubella virus and some plant furoviruses, in particular beet necrotic yellow vein virus (BNYVV).

The methyltransferase domain is required, along with a guanyltransferase, for capping viral plus strand RNA. HEV also encodes an area termed the 'Y' domain located sequentially downstream of the methyltransferase and found in rubella virus and BNYVV. The putative papain-like protease mediates processing of non-structural genes in rubella and alphaviruses. HEV contains motifs similar to the 'X' domain that is associated with such proteases in positive-stranded RNA animal viruses. Regions with strong similarity to the protease were not found in HEV, however, conserved short stretches of amino acids around the putative catalytic cysteine were present. The putative RNA helicase of HEV contained the seven conserved segments typical of such helicases, with the overall closest match with that from BNYVV. The putative RNA-dependent RNA polymerase of HEV was again similar to those from rubella virus and BNYVV, belonging to the supergroup III of viral RNA polymerases. In addition to the above, HEV also contains a region rich in proline, termed the proline hinge, which is also found in rubella virus. Koonin (Koonin et al., 1992) hypothesised that HEV may have evolved from a common ancestor to rubella virus and BNYVV. Since HEV is not enveloped, the possibility remains that HEV could represent a combination of a rubella-like genome with calicivirus-like genome. Alternatively, HEV may be more closely related to an ancestral form, with rubella virus arising as a result of acquisition of glycoprotein genes.

The full length HEV ORF 2 protein has a predicted 660 amino acids with a mass of 72,000 daltons. The calculated pI of 10.35 for part of the ORF 2 gene product suggests it may associate with RNA and act as a capsid protein. The deduced amino acid sequence of ORF 2 contains three potential N-linked glycosylation sites (Asn-X-Ser/Thr) for the Burmese strain, indicating that the product may be a glycoprotein. In addition, the observed immunoreactivity of ORF 2 protein (see diagnosis section 1.3.9) also suggests a structural role. The ORF 3 product was not expected to be utilised in virion assembly, but the identification of ORF 3 antibodies in convalescent patients' sera (Yarbough et al., 1991) suggests it is expressed.
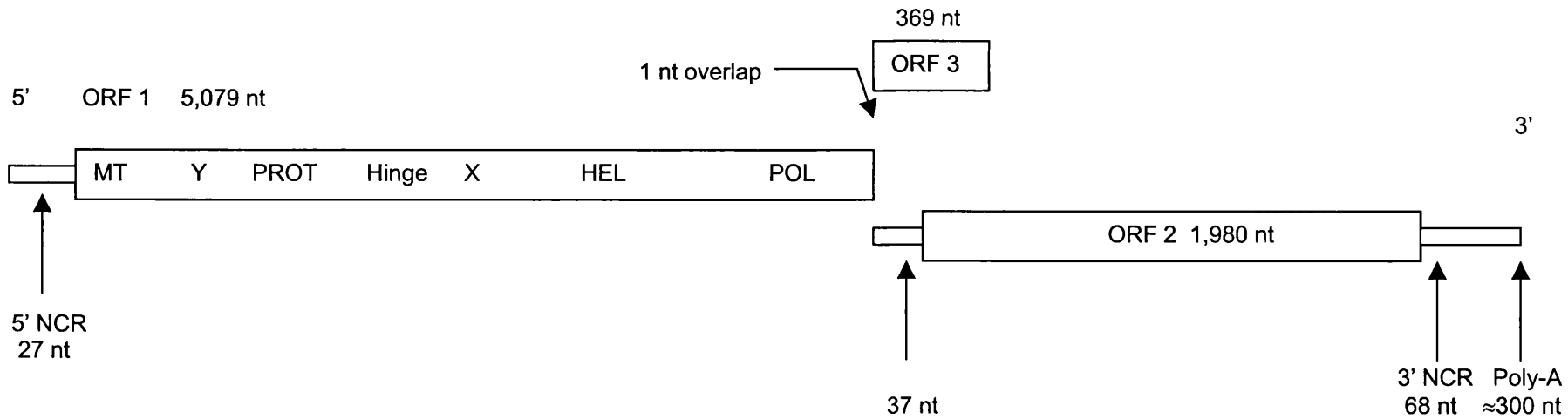
Figure 1.1: Organisation of HEV genome

The schematic representation shows the organisation of the Burmese prototype of HEV with a positive sense RNA of 7,194 nucleotides (nt). NCR- non-coding region; ORF- open reading frame; MT- methyltransferase; Y- domain Y; PROT- papain-like protease; X- domain X; HEL-helicase; POL- RNA dependent RNA polymerase.

ORF 2 translation is in the plus 1 frame from ORF 1.
ORF 3 translation is in the plus 2 frame from ORF 1.
The 5' end of ORF 3 overlaps ORF 1 by one nucleotide.

Comparison of full length, and partial length, HEV sequences from different geographical areas has led to the description of several genotypes based upon restriction enzyme digestion patterns (Gouvea *et al.,* 1998) or sequence identity (Erker *et al.,* 1999; Schlauder and Mushahwar, 2001). The first two isolates of HEV from Burma and Mexico showed an overall nucleotide identity of around 76%. Full length sequences from Pakistan, India, Burma and China often show greater than 90% identity to the Burmese prototype. Many nucleotide differences between strains are at the third base in codon and are therefore silent. Most of the sequence divergence within ORF 1 is located within the hypervariable region. The ORF 2 predicted protein product is relatively conserved across isolates, with an amino acid similarity of around 93% between the divergent Burmese and Mexican isolates. More recently reported isolates from the USA (Erker *et al.,* 1999) and China (Wang *et al.,* 1999) appear to be unique. Identification of a novel HEV in Nigeria (Buisson *et al.,* 2000) most closely related to the Mexican strain, suggests the co-existence of the Burmese and Mexican genotypes in West Africa. A comparison of the nucleotide and amino acid similarities for different fully sequenced HEV isolates is given in the results section. Following the identification of novel HEV isolates in Greece, Italy (Schlauder *et al.,* 1999), Argentina (Schlauder *et al.,* 2000), and Thailand (Kabrane-Lazizi *et al.,* 2001) the existence of eight genotypes was proposed (Schlauder *et al.,* 2000). Genotype 1 comprised Southeast and Central Asian isolates (Burma, Pakistan, China and India); genotype 2 represented by a single Mexican isolate; genotype 3 comprised the USA isolates and possibly one from Thailand; genotype 4 comprised some Chinese/Taiwanese isolates; genotype 5 included an Italian isolate; genotypes 6 and 7 included Greek isolates; and genotype 8 was represented by an Argentinian isolate (see appendix 11 for the table of potential genotype designations). The diversity between the Greek isolates may represent true genetic heterogeneity from native viruses or indicate that the infections may have been acquired from geographically distinct areas. Genotypes 5 through 8 were assigned on partial nucleic acid sequences only, and the possibility exists that comparison of sequences from alternative genomic regions would result in altered similarities. Further division of all HEVs into nine groups has also been proposed (Schlauder and Mushahwar, 2001), taking into account additonal data. This resulted in groupings similar to the genotypes described above, with group 1 including the Burmese prototype and related strains from Asia and Africa; group 2 including the Mexican strain and some from Nigeria; group 3

comprising the US human isolates; group 4 containing the Italian isolate; group 5 containing one Greek isolate and two Spanish isolates; group 6 including the second Greek isolate; group 7 containing Argentine and Austrian isolates; and single unique isolates from China accounting for groups 8 and 9. The swine HEVs from the US (Meng *et al.*, 1997) fall within group 3; a New Zealand swine HEV (Garkavenko *et al.*, 2000) falls into group 4.

A simpler method of classification can result from reducing the number of groupings and allowing more diversity within each genotype (appendix 11, scheme 1). For example, genotype 1 would exhibit little diversity and include strains related to the Burmese prototype, whereas the other 3 genotypes would contain diverse strains: genotype 2 would include the Mexican and Nigerian strains, genotype 3 would include the US and European strains, and genotype 4 would include Chinese/ Taiwanese strains.

Several different phylogenetic analyses have been made for HEV strains, and as new data are included it has become clear that genetically diverse HEVs may co-exist in similar regions. The impact this finding has upon the understanding of HEV hepatitis or vaccine development is yet to be determined.

1.3.4 Hepatitis E virus replication

The replication and transcription strategy of HEV remains poorly understood. The available data would fit a mode involving initial translation of genomic RNA to produce enzymes allowing generation of a full length antigenomic RNA, which is then used to produce further full length genomes and possibly subgenomic RNAs. The existence of subgenomic RNA was suggested by demonstrating hybridisation of a 3' located cDNA to RNA sized at 2.0 kb and 3.7 kb on a northern blot of infected cynomolgus macaque liver (Tam *et al.*, 1991; Yarbough *et al.*, 1991). Internal initiation or differential half-life may explain the higher molar amount of the 2.0 kb species over the 3.7 kb species noted in that study. The two candidate subgenomic RNAs were polyadenylated and therefore were expected to be co-terminal with the 3' end of the genome (similar to that seen in the replication of coronaviruses). However, the detection of subgenomic HEV RNA has not been confirmed and is contentious. Splicing of transcripts has not been found (Tam *et al.*, 1991). A region similar to an internal RNA transcription initiation site of Sindbis virus (Simmons & Strauss, 1972)

has been found in HEV and this maps close to the 5' end of the apparent subgenomic message.

The development of an efficient cell culture system for HEV and the expression *in-vitro* of full length ORFs would be useful in further elucidating the replication mechanism of HEV. Early experiments achieved some success when FRhK-4 cells were co-cultivated with primary kidney cells from an infected cynomolgus macaque (Kazachkov *et al.*, 1992). The development of a serum free medium which could maintain differentiated hepatocytes in long term culture (>100 days) allowed work showing prolonged replication in hepatocytes isolated from infected cynos (Tam *et al.*, 1996). Continual replication was shown by detection of positive and negative strand HEV RNA in cells, but only genomic RNA in culture medium, and by the visualisation of VLPs in concentrated culture fluid by IEM. These VLPs were of unusually small size, 22-32 nm, and the ORF 2 product could not be detected in them by immunoblotting. The size discrepancy may have been the result of concentration and formalin treatment and the immunoblot used may have been suboptimal. No cytopathic effect was seen in the above study, however, a Chinese HEV (strain 87A) originally isolated in human diploid embryonic lung cells, and another strain (93G) showed marked CPE as well as HEV RNA production when cultured in A549 cells (human lung carcinoma) (Huang *et al.*, 1995; Wei *et al.*, 2000).

Currently, the data on expression and processing of the HEV ORFs is derived from experiments using subgenomic fragments of HEV cDNA or *in-vitro* synthesised full length HEV RNAs, expressed in various animal cells or cell-free translation systems. The ORF 2 encoded protein (pORF2) is the putative capsid protein with a predicted size of 660 amino acids and a mass of 72 kDa (Jameel *et al.*, 1996). The N-terminal domain has a high isoelectric point consistent with being involved in genome packaging. Expression in cell culture has produced a variety of products that could be explained by post-translational modification. The amino-terminal 22 amino acids comprise positively charged arginine residues, a hydrophobic core and turn inducing proline residues, which may be a signal sequence directing it to the endoplasmic reticulum, where cleavage and glycosylation may occur (Jameel *et al.*, 1996; Zafrullah *et al.*, 1999; Torresi *et al.*, 1999). Deletion of the putative signal sequence followed by expression in COS-1 cells led to production of a non-glycosylated protein only (Zafrullah *et al.*, 1999). Transit of the ORF 2 protein to the endoplasmic reticulum appears necessary for cell surface expression, whereas glycosylation does not.

Cleavage of the initial 111 amino acids of the ORF 2 protein has been observed following expression in insect cells of recombinant ORF 2 (Zhang *et al.*, 1997; Robinson *et al.*, 1998). Glycosylation may occur at any one of three N-linked sites (Asn 137, Asn 310, Asn 562; motif Asn-X-Thr/Ser), typically resulting in an 88 kDa protein expressed intracellularly and on the cell surface (Jameel *et al.*, 1996; Zafrullah *et al.*, 1999). Since the capsid protein of HEV has not been characterised, the role of glycosylation in replication within hepatocytes is uncertain, however it may affect protein folding or influence the polarity of cellular secretion. Although most of the ORF 2 protein produced by heterologous expression is glycosylated, this may be an artefact. Expression of ORF 2 in mammalian cells (HepG2) resulted in the production of 78 kDa, 82 kDa, and 86 kDa proteins, with the larger two representing glycosylated forms (Torresi *et al.*, 1999). The glycosylated forms of the ORF 2 protein accumulated at the cell surface, and were not stable in the HepG2 cells, whereas the non-glycosylated form accumulated in the cytosol, and may have functioned as the capsid precursor protein. Pulse chase experiments expressing ORF 2 alone in an SV40 based vector and COS-1 cells, resulted in a non-glycosylated precursor protein of 82 kDa. It is unclear what effect the presence of the other HEV proteins and RNA have on ORF 2 protein processing during natural infection. Glycosylation could be blocked or the glycosylated ORF 2 product may be stabilised. Further evidence that the non-glycosylated ORF 2 protein is the capsid precursor comes from expression in insect cells of an ORF 2 gene lacking the signal sequence (deletion of initial 111 amino acids), which led to self-assembly of the 50 kDa products into virus-like particles (Li *et al.*, 1997). Expression of capsid genes of other caliciviruses has resulted in formation of empty virions or VLPs (Norwalk virus- Jiang *et al.*, 1992 and Rabbit Haemorrhagic Disease Virus- Laurent *et al.*, 1994). Furthermore, if the capsid protein of HEV was glycosylated, the virions might be expected to have an envelope.

The ORF 3 product is around 13.5 kDa, is not glycosylated (Jameel *et al.*, 1996; Panda *et al.*, 2000), and was not expected to be utilised in virion assembly, particularly as pORF2 alone can form virus-like particles. The identification of an antibody response in infected patients (Yarbough *et al.*, 1991) suggests it is expressed. It has two hydrophobic N-terminal domains, which may be transmembrane, and it appears to be phosphorylated (Zafrullah *et al.*, 1997). ORF 3 subgenomic message has not been found, although expression could result from alternate initiation codon usage of a single subgenomic for both ORF 2 and 3. Interaction with a cellular protein has

been suggested, since immunoprecipitation of ORF3 protein led to coprecipitation of an 18kDa protein if expression was in COS-1 or Hu7 hepatoma cells, however this was not observed with HepG2 cells (Jameel *et al.*, 1996). Confirmation came from precipitating this cellular protein from mixtures of untransformed COS-1 cell lysates and radiolabelled pORF3 using anti-pORF3. The role of this cellular protein is uncertain. The differences between the cell lines could result from the different stages of differentiation.

The entire ORF 1 has been expressed in *E. coli*, cell free translation systems and in various mammalian cells. ORF 1 seems to encode a protein (pORF1) of around 186 kDa when expressed alone (Ropp *et al.*, 2000; Ansari *et al.*, 2000), corresponding to the size predicted by nucleotide translation. The non-structural proteins thought to be encoded by ORF 1 include a papain-like cysteine protease, and therefore post-translational cleavage of the full length ORF 1 polyprotein might be expected. Characterisation of the ORF 1 protein by western blotting and immunoprecipitation following expression in a cell free mammalian translation system and HepG2 cells, showed a protein of around 186 kDa (Ansari *et al.*, 2000). No processing of the ORF 1 product was seen, despite incubation at 37°C for 24 hours or incubation in the presence of different metal ions. Similar results were obtained by another group using an *in-vitro* coupled transcription system; processing of the 185 kDa protein was not observed despite addition of divalent cations and microsomes (Ropp *et al.*, 2000). However, expression *in-vivo* in conjunction with extended incubation (24-36 hours) yielded two potential processed products of 107 kDa (N-terminal) and 78 kDa (C-terminal). Mutagenesis of one of the predicted catalytic sites (Cys 483) within the proposed protease failed to abolish the observed cleavage, and the other site (His 590) was not conserved. Therefore, the observed cleavage could not be attributed to the presence of an ORF 1 encoded protease.

Recently, transfection of HepG2 cells with an *in-vitro* synthesised RNA from a full length HEV cDNA has provided additional insight into the replication of HEV (Panda *et al.*, 2000). The HEV clone constructed differed from the wild type template by the presence of one nonviral nucleotide at the 3' end, 12 nonviral nucleotides at the 5' end, and the absence of a 5' cap. HEV replication was demonstrated by the detection of negative sense HEV RNA and the production of HEV infection in a rhesus monkey, following inoculation with cell culture supernatant. Viral proteins were

detected by immunoprecipitation and autoradiography. The following proteins were identified: 72 kDa ORF 2, 13.5 kDa ORF 3, 35 kDa corresponding to the putative methyltransferase, 38 kDa helicase, and 36 kDa RDRP. Cell sampling at different times showed the putative helicase and methyltransferase proteins were present at 72 hours post-transfection, whereas the putative RDRP was present at 12, 24 and 36 hours post-transfection, but not at 72 or 96 hours. This suggested that the RDRP is an early protein, which undergoes rapid degradation. Processing of ORF 1 proteins may only occur in the context of complete genome expression. The putative protease of HEV has not been demonstrated, and therefore other viral proteins within the cellular environment may activate alternative proteases.

Recent characterization of HEV genotype 4 from China (Wang et al., 2000) showed a single base insertion at position 5159 which would add 14 residues to the 5' end of the ORF 2 product, when compared to other geographically distinct isolates. These extra amino acids would extend the signal sequence, although cleavage by the signal peptidase may result in a standard length ORF 2 protein. This base insertion is also likely to result in a shorter ORF 3 protein of 112 amino acids, if translated from the first possible codon. The consequences of these changes on replication efficacy and pathogenesis are not clear.

1.3.5 Epidemiology of hepatitis E in endemic and epidemic areas

The development of diagnostic assays for HEV infection has shown hepatitis E to be a disease of widespread distribution, however, its greatest impact is in resource poor countries where it is believed to account for greater than 50% of acute viral hepatitis in young to middle-aged adults. Estimates implicate HEV as the cause of around 2 million cases of hepatitis in India per year where it is a common cause of jaundice (Datta et al., 1987) and acute liver failure (Acharya et al., 1996). Epidemics have been documented in many areas including Pakistan (Iqbal et al., 1989), India (Tandon et al., 1982; Singh et al., 1998), South East Asia (Burma, Myint et al., 1985), Nepal (Kane et al., 1984), Africa (Belabbes et al., 1985; Tsega et al., 1991; Mushahwar et al., 1993) and North America (Mexico, Velazquez et al., 1990). Some of these areas experience recurrent epidemics. The largest documented outbreak struck the Xinjiang Uighar region of North West China between 1986-88, with over 119,000 cases of jaundice (Zhuang et al., 1991).

No outbreaks of hepatitis E have been recorded in the USA, Canada, Europe or developed areas of Asia. Some countries, such as Egypt, Turkey and Hong Kong appear to have a significant number of sporadic cases of hepatitis E yet have never reported an epidemic (Goldsmith *et al.*, 1992; el-Zamaity *et al.*, 1993; Coursaget *et al.*, 1993; Lok *et al.*, 1992).

Hepatitis E is spread by the faecal-oral route and occurs in large and small epidemics and sporadically. These epidemics can be unimodal or prolonged with several peaks (Naik *et al.*, 1992) and most are associated with gross faecal contamination of the drinking water supply during the rainy season; adequate chemical treatment of such water often causes the epidemic to subside. The first extensively studied epidemic (originally attributed to HAV, but this was serologically excluded later by Wong *et al.*, 1980) occurred in Delhi, North India in 1955-6 (Viswanathan, 1957) when sewage contamination of the Yamuna river led to 29,000 cases of jaundice. An association between the high prevalence of anti-HEV and the use of river water for washing, drinking and disposal of excreta has been found for parts of South-East Asia (Corwin *et al.*, 1999).

Person to person transmission rates are low at only 1-2% (Myint *et al.*, 1985; Aggarwal *et al.*, 1994; Singh *et al.*, 1998) in the setting of a general population incidence of around 5%. This is in contrast to HAV which has a secondary household attack rate of over 20%. This low rate of household transmission may be linked to a need for a high viral dose likely only to come from grossly contaminated drinking water. The difference between rates of inter-personal spread for HAV and HEV may also be the result of the relatively high environmental stability of HAV. The role of intrafamilial spread was investigated in detail in a large bimodal epidemic in Kanpur, India (Aggarwal *et al.*, 1994). The majority of secondary cases occurred before the low limit of the incubation period for HEV (two weeks). In addition, new cases stopped appearing around eight weeks after purification of the water supply. Previously studied epidemics with multiple peaks of incidence are probably best attributed to repeated contamination of the water supply.

The vast majority of cases of hepatitis E result from faecal-oral spread but the possibility of parenteral spread has been investigated (discussed in section 1.3.7).

Seroprevalence studies of HEV antibody have been conducted on a geographical basis and on specific demographic groups such as haemodialysis patients, haemophiliacs, male homosexuals, HIV infected patients, and patients with other forms of liver

disease. In general, HEV IgG seroprevalence data illustrate the link with relatively underdeveloped areas and use of untreated drinking water, correlating with the main mode of transmission. This was supported by finding a rate of 5.8% for HEV antibody in urban areas of South Africa and 19% in the rural areas (Tucker *et al.*, 1996). In contrast, a survey in Hong Kong, where drinking water is chlorinated, found HEV IgM in 16.5% cases of acute viral hepatitis and HEV IgG in 16% of healthy subjects (Lok *et al.*, 1992), suggesting a high background rate of infection rather than recently imported disease from nearby countries.

Early studies will have been influenced by the variable sensitivities and specificities of the antibody assays used, however, it appears that in endemic areas, anti-HEV IgG is found in up to 5% of children under the age of 10 years and in around 10-40% of adults (Lok *et al.*, 1992; Thomas *et al.*, 1993; Mushahwar *et al.*, 1993; Arif *et al.*, 1994; Paul *et al.*, 1994; Arankalle *et al.*, 1995a; Aubry *et al.*, 1997; Clayson *et al.*, 1997; Alvarez-Munoz *et al.*, 1999). Many studies show a sharp increase in antibody prevalence between childhood and adolescence/ adulthood. Seroprevalence has been determined in populations from Pune in India (Arankalle *et al.*, 1995) using samples from 1982 and 1992. The authors found that although most individuals had been infected with HAV by age 3 years, the peak prevalence of anti-HEV (around 40%) was in young adulthood. The low rate of HEV antibody in children less than 10 years old (true for both years of study) is interesting and has been confirmed in other studies (Lok *et al.*, 1992; Mushahwar *et al.*, 1993; Thomas *et al.*, 1993; Darwish *et al.*, 1996). The changing incidence with age could not be explained by a gradual accumulation in exposure as there was a sharp increase between childhood and adulthood. The age related prevalence of HEV antibody has not been fully explained, but may be partially accounted for by several factors (see below).

During an outbreak, the clinical case rate is around 5% but can vary between 1 and 15% (Khuroo, 1980; Kane *et al.*, 1984; Tsega *et al.*, 1992; Naik *et al.*, 1992; Singh *et al.*, 1998) with a higher relative proportion of young to middle aged adult cases. A striking feature of many of the epidemics of HEV infection is the high incidence of severe disease in the later stages of pregnancy (see section 1.3.12). The data on the separate attack rates for men and women are contradictory (see section 1.3.12). The age distribution of cases is consistent with the observation that in endemic areas the prevalence of antibody sharply increases around adulthood. Explanations for the reduced seroprevalence and relatively low number of symptomatic cases in children

include lower exposure, resistance to infection, rapid loss of antibody and non-production of antibody or a protective immune response following infection.

There is no evidence for inherent resistance to HEV disease in children; symptomatic hepatitis E has been observed in them (Hyams *et al.*, 1992; Goldsmith *et al.*, 1992; Khuroo *et al.*, 1994; Malathi *et al.*, 1998; Arora *et al.*, 1999). Decreased exposure of children could be the result of living through a period of low HEV activity, a relative lack of travel to new areas (necessary in adults for work) and ingestion of smaller volumes of food and water. It should be noted that these conditions do not affect the high incidence of HAV infection in children. It is possible that antibody is lost rapidly after initial infection at an early age and the re-exposure rate determines when persistent, detectable antibody levels develop. It is interesting to note that gastroenteritis due to Norwalk virus, a calicivirus, may not induce long-term protective immunity. Initial infections could also be more likely to be subclinical, with an incomplete immune response, with re-infection resulting in worse disease as a result of an enhanced immune response. A similar model has been proposed for Dengue shock syndrome (Halstead, 1988), where the presence of pre-existing antibody (maternal in infants or acquired from previous infection in adults) against one serotype leads to a pathological increased activation of cellular immunity during infection by a different serotype. Alternatively, an inadequate immune response to the first HEV infection may allow re-infection with the same serotype and result in an enhanced immune mediated cell destruction. This may be the mechanism involved in the more severe natural RSV infection seen in children previously immunised with a formalin inactivated RSV vaccine (reviewed in Domachowske and Rosenberg, 1999). The vaccine may not have led to protective levels of locally produced IgA antibody in the lungs but the cellular immune response had been primed.

The existence of subclinical hepatitis E was originally suggested by an excess of anti-HEV antibody prevalence above the reported history, or observed frequency, of jaundice in both endemic and non-endemic areas. Further evidence followed, including the documentation of seroconversion for IgM and IgG HEV antibody in a Dutch woman without concomitant illness (Zaaijer *et al.*, 1993). A study in Nepal (Clayson *et al.*, 1997) found over an 18 month period that subclinical infection with HEV (seroconversion only) was more common than infection resulting in jaundice (54 versus 18 cases respectively). Although the history of jaundice was based upon the subjects' recollections alone, and other symptoms could have been present at the

time of infection, this finding strongly supported the existence of subclinical hepatitis E. The study population of 757 individuals contained only seven women and no children below the age of 12, therefore it is not certain whether the results are directly applicable to the general local population and other areas.

It is unclear whether the longevity of the HEV antibody response can explain some of the unusual epidemiological features of hepatitis E. There are conflicting data regarding the persistence and production of HEV antibody, but such studies will have been influenced by assay design and sensitivity (see diagnosis section 1.3.9). Non-production of HEV antibody after infection has been observed (Ke *et al.*, 1996, Clayson *et al.*, 1995b). Disappearance of IgG antibody has been noted as early as 12 months (in children- Goldsmith *et al.*, 1992, in adults- Skidmore *et al.*, 1992) or may last much longer (Khuroo *et al.*, 1993; Lee *et al.*, 1994; Chadha *et al.*, 1999). In an isolated case from Austria, HEV IgG was no longer detectable using a commercial assay 56 days after hospitalisation (Worm *et al.*, 1998) whereas in a volunteer study, IgG was still detectable over 4 years after the infection (Chauhan *et al.*, 1998). The longevity of antibody following an outbreak of hepatitis E in Kashmir in 1978 has been studied (Khuroo *et al.*, 1993). In 1992 sera from 45 patients affected in the 1978 outbreak were retested for anti-HEV and nearly half were positive. This finding was unlikely to represent reinfection (icteric or anicteric) as none of the patients reported jaundice in the intervening period and a control population showed no rise in the low IgG prevalence over the test period. The effect of variable assay performance on these studies has not been quantified but may have been significant. In a study of HEV antibody prevalence in San Marino (Rapicetta *et al.*, 1999), a significant proportion of subjects (8 of 22) who had positive HEV EIA results confirmed by western blot, had not maintained detectable levels of HEV antibody after five years. The predictor of persisting antibody was a sample to cut-off ratio of over two. Despite these findings, the reason for the observed seroprevalence of anti-HEV IgG across age groups remains unexplained.

1.3.6 Epidemiology of hepatitis E in 'non-endemic' areas

A growing number of sporadic cases of HEV infection have been documented in developed countries with no clear evidence of outbreaks or endemicity, and the great majority of these are imported from endemic areas. However, some of these cases are without obvious risk factors (Italy, Zanetti *et al.*, 1994 and 1999; Netherlands, Zaaijer *et al.*, 1995; Germany, Wang *et al.*, 1993; Spain, Buti *et al.*, 1995; Greece, Tassopoulos *et al.*, 1994; Austria, Worm *et al.*, 1998; USA, Kwo *et al.*, 1997; United Kingdom, McCrudden *et al.*, 2000), suggesting HEV infection should be considered even in the absence of a relevant travel or contact history. The diagnosis was often made by serological exclusion of other agents plus detection of HEV antibody (IgM, or IgG seroconversion or both). It should be remembered that such findings do not necessarily indicate a background of typical HEV hepatitis. Alternative explanations could include contact with a subclinical or unreported imported case, ingestion of contaminated foodstuffs, infection with an atypical HEV circulating within Europe which is less pathogenic, or contact with an unidentified animal reservoir (for example, pigs). Serological analysis of possible locally acquired cases of HEV infection outside accepted endemic areas may underestimate the true proportion due to HEV. In a reported HEV infection thought to have been acquired in North America, standard HEV IgM antibody assays were negative, although IgG was detected (Kwo *et al.*, 1997). This HEV isolate was later shown by sequence analysis to be divergent from others. In addition, anti-HEV was not detected in serum samples taken from individuals infected with genotype 4 HEV using assays based upon Burmese/ Mexican antigens (Wang *et al.*, 2000).

Infectious (for rhesus monkeys) HEV has been detected in raw sewage in Spain (Pina *et al.*, 1998). This first report showed the nucleotide sequence was most similar to an Indian HEV, which may have been imported, however, viable virus was clearly recovered from the environment. Further studies confirmed the presence of HEV RNA in pig serum, faeces and pig slaughterhouse sewage samples from Spain (Pina *et al.* 2000). The RNA found in the sewage samples showed 92-94% nucleotide similarity to those found in serum taken from Spanish patients with HEV hepatitis. The discovery of a sequence from variant HEVs endogenous to Italy (Zanetti *et al.*, 1999) and Greece (Tassopoulos *et al.*, 1994) has strengthened the case for the existence of this agent in Europe. The Italian isolate was found as part of a study of

HEV in Italy, in which several patients had no travel risk factors. Nucleotide sequencing of a short segment (around 300 nucleotides) revealed the closest homology was with an isolate from the USA (Schlauder *et al.,* 1999).

In addition to these cases, a small percentage of the healthy indiginous populations from areas believed to be non-endemic show HEV IgG seroreactivity. Cases of HEV infection might be expected in the Mediterranean region given its proximity to Africa and the Middle East, but this finding extends to other areas also. Studies to determine the background rate of HEV antibody have concentrated upon the blood donor population. Blood donors within Europe have a background rate of HEV IgG antibody of around 1-5% (Dawson *et al.,* 1992; Zaaijer *et al.,* 1992 &1993; Medrano *et al.,* 1994; Zanetti *et al.,* 1994; Paul *et al.,* 1994; Mateos *et al.,* 1999). Similar rates exist in healthy subjects in Israel (Karetnyi *et al.,* 1995). In Australia (Moaven *et al.,* 1995), HEV IgG was found at the slightly lower rate of 0.4% in blood donors. In the USA a similar low level prevalence of antibody has generally been found (Mast *et al.,* 1997) with the exception of one study which reported an HEV IgG prevalence of 21% in blood donors from Baltimore (Thomas *et al.,* 1997), 23% in intravenous drug users and 15% in homosexual men. No behavioural factors were identified which could account for this high seroprevalence, which remains unexplained.

A survey of patients from inner-London attending their general practitioners for conditions unrelated to hepatitis, found HEV IgG in 3.9% of UK born subjects and 8.8% of those not born in the UK (Bernal *et al.,* 1996). The age specific analysis suggested a cohort effect for those born in the UK, possibly linked to enhanced prevalence in the past, whereas a higher rate of seroprevalence was found in the younger non-UK born subjects.

In all of these studies the detection of antibody could be explained in several ways. Long-lived antibody from distant infections, sub-clinical or unreported disease due to a typical or variant HEV or poor specificity of the relevant assays. It should be noted that in many of these studies involving low risk groups in non-endemic areas, the reactivity in the HEV IgG assays often has a low sample to cut-off ratio in comparison to samples generated during epidemics.

1.3.7 Parenteral transmission

The vast majority of cases of hepatitis E result from faecal-oral spread, but other modes have been the subject of research. Parenteral transmission of HEV is theoretically possible via transfusion, as there is a viraemia prior to the onset of symptoms. Hepatitis E virus RNA has been detected by RT-PCR in 3 of 200 Indian blood donors at the time of donation (Arankalle and Chobe, 1999). None of these donors showed symptoms of HEV infection around donation or on follow up. Unlike for HAV, such transfusion transmission of HEV has not yet been clearly demonstrated. A retrospective serological analysis of transfusion recipients in India indicated that a significantly higher proportion of susceptible transfusion recipients developed IgM anti-HEV during a 2 month follow up period when compared to susceptible donors (2 of 37 versus none of 34) (Arankalle and Chobe, 2000). Neither of the 2 patients developed symptoms of hepatitis, however, one showed a rise in biochemical marker of liver damage.

Indirect evidence for transfusion transmission has been sought by investigating the seroprevalence of HEV IgG in recipients of blood products (In Italy- Mannucci *et al.*, 1994; the Netherlands- Zaaijer *et al.*, 1995; in Israel- Barzilai *et al.*, 1995; in Germany- Wang *et al.*, 1993). These studies were limited by the use of blood products coming from donors living in areas without HEV epidemics and the possibility of other confounding factors such as chronic hepatitis, and a lack of travel history information.

The parenteral route for HCV and HBV transmission is well recognised, and a higher seroprevalence for HEV IgG has been demonstrated in those individuals infected by HCV or HBV (Italy- Pisanti *et al.*, 1994; and France- Halfen *et al.*, 1994). It was unclear whether this reflected true parenteral transmission, the presence of false positive assay reactivity in the presence of chronic hepatitis, or a cohort effect from subjects living through a period when HEV was more common. In contrast, HEV antibody was not found to be associated with other blood-borne viruses in haemodialysis patients, from Dublin (Courtney *et al.*, 1994), Sweden (Sylvan *et al.*, 1998) or Italy (Fabrizi *et al.*, 1997). A higher rate of seropositivity has been observed amongst intravenous drug abusers in some countries (Switzerland- Lavanchy *et al.*, 1994; Italy- Zanetti *et al.*, 1994).

Parenteral spread of HEV was also suggested as a possible mode when three health care workers developed hepatitis following caring for a pregnant woman in South Africa (Robson *et al.*, 1992). However, only one of the staff members had serological evidence of HEV infection and contact with the patient's faeces had also occurred.

There is no evidence for sexual spread (Myint *et al.*, 1985; Aggarwal *et al.*, 1994), however a link between HIV and HEV transmission has been proposed (Balayan *et al.*, 1997) but remains unconfirmed, with at least one study in Africa showing no link (Aubry *et al.*, 1997).

Vertical transmission has been documented (Khuroo *et al.*, 1995) in humans (see section 1.3.13), but there is no clear evidence for this in experimentally infected primates (Arankalle *et al.*, 1993; Tsarev *et al.*, 1995).

1.3.8 Hepatitis E as a zoonosis

Despite the existence of significant numbers of sporadic HEV infections, it is unclear whether HEV is maintained between epidemics simply by horizontal serial transmission between humans or by other factors. The persistence of HEV in the environment outside of water systems is possible but has not been shown. Many investigations have studied animals kept in close proximity to humans, since these could be responsible for sporadic disease and prolonging epidemics. HEV has an experimental host range including rats (Maneerat *et al.*, 1996), pigs (Balayan *et al.*, 1990; Meng *et al.*, 1998a) and sheep (Usmanov *et al.*, 1994), and efforts have been made to identify natural infection within those animals. HEV antibodies have been found in wild caught monkeys (Ticehurst *et al.*, 1992b; Arankalle *et al.*, 1994b; Tsarev *et al.*, 1995), suggesting prior exposure to HEV or an antigenically related agent.

HEV antibody and HEV RNA have been detected in a proportion of free roaming domestic pigs in Nepal (Clayson *et al.*, 1995) and this was followed by characterisation of a specific pig HEV in the USA (Meng *et al.*, 1997). By the age of 3 months, most pigs in mid-Western USA are seropositive for antibody cross-reacting with human anti-HEV (Meng *et al.*, 1997), and there is evidence of HEV infection in pigs in other countries where HEV is uncommon (Spain- Pina *et al.*, 2000; New Zealand- Garkavenko *et al.*, 2000) and where HEV is common (China and Thailand- Meng *et al.*, 1999; Taiwan- Hsieh *et al.*, 1999; India- Arankalle *et al.*, 2001).

Clinically apparent illness in pigs associated with swine HEV infection has not been observed. Whereas pigs do not appear to be susceptible to the Pakistan or Mexico isolates of human HEV (Meng *et al.*, 1998a), or an Indian strain (Arankalle et al., 2001), they do appear to be susceptible to infection with a USA strain (US-2) of HEV, which shares over 97% amino acid identity with the USA swine HEV (Meng *et al.*, 1998c). This USA sourced swine HEV could also infect rhesus monkeys and chimpanzees. A later study showed that pigs experimentally infected (by intravenous inoculation) with the human US-2 strain of HEV suffered more severe and persistent hepatic lesions than those animals inoculated with a swine HEV from the USA, however no pigs developed symptoms or biochemical markers of deranged liver function (Halbur *et al.*, 2001). The presence of HEV in infected pig faeces and liver tissue suggests there may be a risk of transmission of swine HEV to individuals in close contact with pigs or to human recipients of pig liver/ hepatocyte xenografts.

A prolonged viraemia has been noted in some pigs (Meng *et al.*, 1998a), making them an important candidate reservoir, particularly in areas where non-human primates are absent. A possible link with environmental HEV exposure was suggested by the detection of a significantly higher prevalence of anti-HEV in healthy field workers in Iowa, USA in comparison to normal blood donors from the same area (Karetnyi *et al.*, 1999). Comparison of the prevalence of anti-HEV in Taiwanese pig handlers and matched controls showed a higher rate in the pig handlers, however this was only marginally significant (Hsieh *et al.*, 1999). In addition, studies in Taiwan have shown that local swine HEV and human HEV (individuals with no travel risk factors outside Taiwan) form a monophylogenetic group based on nucleotide identity (Wu *et al.*, 2000). Subclinical human infections with Taiwanese swine HEV may account for the relatively high prevalence of anti-HEV in the general population.

Antibody which reacts in an HEV IgG assay has been found in a high proportion of wild rats caught in parts of the USA; in 90% from Hawaii, 77% Maryland and 44% Louisiana (Kabrane-Lazizi *et al.*, 1999). Furthermore, the prevalence of antibody appeared to increase with the estimated age of the rats, implicating them in the maintenance of HEV in the environment. Other investigators have found a high proportion of rodents in the USA to be positive for anti-HEV, with the antibody prevalence higher in urban caught animals than in the rural caught animals (Favorov *et al.*, 2000). The highest antibody prevalence was 59.7% in the genus *Rattus*, possibly reflecting their high population density. Notably, some seropositive rodents

were found in remote areas, away from regular human contact. In addition, anti-HEV IgG has been detected in up to 6.9% of cattle, up to 21.5% rodents, and 22.7% dogs in one Indian study (Arankalle *et al.*, 2001), however anti-HEV was not found in goats. There is no direct virological or epidemiological evidence that HEV is a zoonosis, however it is clear that there is potential for cross-species exposure, and whether this leads to infection, with or without illness, may be HEV strain dependent.

1.3.9 Laboratory diagnosis of hepatitis E infection

Laboratory diagnosis of HEV infection initially involved IEM of stool samples or a fluorescent antibody blocking assay. These methods were specific but had limited sensitivity, and the development of useful antibody assays and genome detection followed the cloning and sequencing of prototype HEV strains. Tests for IgG, IgM and IgA have been designed. Formats using recombinant antigens, peptides, mosaic proteins comprising several linked immunoreactive peptides, and mixtures of these three have been described (Dawson *et al.*, 1992b; Chau *et al.*, 1993; Tsarev *et al.*, 1993b; Favorov *et al.*, 1994 & 1996; Li *et al.*, 1994; He *et al.*, 1995).

There are few data comparing the sensitivity and specificity of the various antibody assays developed. Considerable differences have been observed between assays when tested against panels of sera (Mast *et al.*, 1998; Ghabrah *et al.*, 1998) and when different assay formats or component antigens were used to confirm initially positive results within the same study (Vandenvelde *et al.*, 1994; Buisson *et al.*, 1994; Quiroga *et al.*, 1996; Zanetti *et al.*, 1994; Zaaijer *et al.*, 1993; Coursaget *et al.*, 1993a). Such assay variation may be more important when evaluating populations non-endemic for HEV. Therefore, early epidemiological studies may have been complicated by the use of assays that were sub-optimal.

Several studies have identified immunoreactive epitopes throughout the HEV ORFs. These include the RNA dependent RNA polymerase in ORF 1, the carboxyl terminal end of ORF 3 and at least six regions in ORF 2 (Kaur *et al.*, 1992; Khudyakov *et al.*, 1994 and 1999). A recent study used murine monoclonal antibodies, generated against the C-terminal 267 amino acids of the ORF 2 product, to investigate the antigenic structure of part of the capsid protein (Riddell *et al.*, 2000). Cross-inhibition in competitive ELISAs was observed between monoclonal antibodies directed against widely separated epitopes, suggesting a complex antigenic structure. Monoclonal

antibodies specific for the conformational epitopes blocked the antibody reactivity of HEV convalescent patient sera, indicating that conformational epitopes may be important in recovery from infection. However, the modelling of some antigenic regions has given inconsistent results depending upon the length of the synthetic peptide or recombinant protein used.

Despite the genomic variability of HEV, several geographically distant isolates have at least one major cross-reactive epitope, encoded by ORF 2, and therefore assays should employ antigens based on this ORF. However, IgM antibody resulting from infection by a novel HEV isolate from the USA was not detectable using antigen derived from Burmese or Mexican HEV (Schlauder et al., 1998). Similarly, anti-HEV was not detected in serum samples taken from individuals infected with genotype 4 HEV using these Burmese/Mexican based antigens (Wang et al., 2000), however it was unclear if sufficient time had elapsed for antibody generation. Assays using only antigens derived from ORF 3 are inadequate (Ghabrah et al., 1998) as they can demonstrate strain specific antigenicity, particularly between the Burmese and Mexican isolates (Yarbough et al., 1991). Commonly employed commercially available HEV antibody assays are supplied by Abbott Laboratories (USA) and Genelabs Diagnostics (USA). These use recombinant antigens derived from the ORF 3 protein and C-terminal end of the ORF 2 protein of the Burmese and Mexican HEV isolates. In order to be comprehensive, future assays may need to incorporate antigens derived from the increasing number of novel HEV isolates. Currently, serological tests may be underestimating the prevalence of HEV across geographical regions.

Serological diagnosis of acute HEV infection may be made by the detection of specific IgM or by the demonstration of seroconversion for IgG. There is a spectrum of IgM responses, possibly limiting the use of IgM in diagnosis of acute HEV infection. It may be detectable as early as 4 days after the onset of illness (Ke et al., 1996), or it may only develop to detectable levels slowly or it may not be detected at all or decline in titre rapidly (Skidmore et al., 1992; Nanda et al., 1995; Clayson et al., 1995). In general, IgM is detected first just prior to the maximal liver injury (peak ALT) and disappears after around 5 months (Favorov et al., 1992), although it has been measured at up to 6 months after jaundice (Lok et al., 1992). Confirmation of recent infection may also be provided by the detection of HEV IgA, although only around 50% of patients' sera show reactivity (Chau et al., 1993). There is also a spectrum of IgG responses, but in general it appears shortly after IgM, and it is then

found at high titre in the majority for between 1 and 4 years (Favorov *et al.*, 1992; Dawson *et al.*, 1992) and possibly for much longer. One study has shown non-production of detectable HEV IgG despite an IgM response (Ke *et al.*, 1996). In this situation a false negative result cannot be excluded without confirmation using different HEV IgG assays.

Serological diagnosis of HEV infection may be complicated by the presence of non-specific reactivity in the assays, which can be more common in patients with viral hepatitis. This is seen in hepatitis A infection where production of non-specific IgM and IgG and rheumatoid factor occurs. Such immune activation by a hepatitis agent unrelated to HEV could lead false positive antibody results.

The best laboratory indicator of HEV infection appears to be detection of viraemia. HEV RNA can be detected in serum and stool using RT-PCR, a typical amplicon being a fragment of the RDRP (Ray *et al.*, 1991), however, appropriate samples and the facilities required are rarely available. Amplification of HEV cDNA also allows genotyping on the basis of restriction endonuclease patterns (Gouvea *et al.*, 1998).

1.3.10 Human hepatitis E infection
Clinical features

Data from epidemics, sporadic cases and volunteer studies have helped elucidate the features of hepatitis E in the human host, however little is known about the pathogenesis of HEV infection in humans. During an epidemic there is around a 5% attack rate in the population at risk, most disease occurring in the 15-40 year age range, with an overall mortality of 0.5-4% (Gust and Purcell, 1987; Bradley, 1992; Krawczynski, 1993), possibly influenced by factors such as local health service provision and the presence of other endemic diseases. Clinically overt infection does occur in children, including some cases of fulminant hepatitis (Arora *et al.*, 1996).

The disease presents like other viral hepatitides with typical biochemical findings and usually runs an acute course with full resolution. Despite this, there is significant mortality, predominantly from fulminant hepatic failure. Subclinical infection almost certainly occurs, but the proportion of cases has not been determined. Following faecal-oral transmission there is an incubation period of between 2-9 weeks with an average of around 40 days. There may be a prodrome 1-7 days prior to the onset of jaundice, consisting of one or more symptoms such as abdominal pain, malaise,

anorexia, nausea and vomiting and mild pyrexia (<38°C). These symptoms often persist during the icteric phase. A study of 423 hospitalized males who had acute hepatitis E (Tsega *et al.*, 1991) found malaise, anorexia, nausea, vomiting, fever and headache to be very common; itching and joint pains were observed in only 13.6% and 7.4% respectively. Abdominal pain was elicited in 82%, but only 9.7% had an enlarged liver on clinical examination; a rash was occasionally seen, as was enlargement of the spleen. A much higher rate of hepatomegaly has been noted elsewhere (Clayson *et al.*, 1995).

The hepatitis is generally self-limiting, resolution usually occurring within 2 weeks (range 1-6 weeks), with no carrier state, but prolonged disease and extended viral shedding has been recorded (see section 1.3.14.2). Fulminant hepatitis occurs sporadically or during epidemics in all patient groups, but particularly in pregnant women in the third trimester (up to 20%) (see section 1.3.12).

Co-infections and super-infections of HEV with other hepatitis viruses occur in countries where they are common and this can influence the course of the disease (see section 1.3.15). The role of human immunodeficiency virus infection, endemic in many of the areas where HEV is common, in the clinical outcome of HEV hepatitis is unclear.

1.3.11   Volunteer studies

The clinical course of disease following experimental volunteer transmission was described first by Balayan in 1983. A later volunteer study (Chauhan *et al.*, 1993) expanded upon this by assessing the biochemical and viral kinetics also. The inoculum containing characteristic 27-34nm VLPs used in the latter study was prepared from stool from an individual with ET-NANB hepatitis acquired in North India. A clarified 10% stool suspension was ingested, thereby mimicking the natural route of infection. Anicteric hepatitis occurred 30 days later accompanied by anorexia, epigastric pain and discoloured urine. Icterus started on day 38 and persisted until day 120, roughly paralleling the rise in alanine aminotransferase (ALT). VLPs which transmitted disease to rhesus monkeys, were seen in the stools by solid phase immune electron microscopy on days 34 and 37. The stools became positive by RT-PCR for HEV RNA at the onset of icterus and remained positive until day 46, approximately at the time of the peak ALT. However, no stool was passed for three days after that time,

and collection was stopped. HEV RNA appeared in serum on day 22 (before disease onset) and persisted until the ALT peaked on day 46 (see appendix 12 for the relationship between viral excretion, presence of icterus and timing of antibody production). These features generally concur with findings in non-human primates. HEV IgG was detected first on day 41 and remained detectable for at least 4 years. Retrospectively tested sera were positive for IgM between 3 and 43 days post-icterus (Nanda et al., 1995), therefore the role of antibody in the initial hepatitis is unclear. The peak ALT was at eight days post-icterus in this volunteer whereas evaluation of peak ALT levels in outbreaks show this can occur earlier, at 0-3 days post-icterus. Notable in this case was the severe nausea and vomiting, the prolonged icterus and slightly shorter incubation period than the average. Two factors may have accounted for this. The inoculum may have been much larger than that usually encountered in natural settings. It has been demonstrated using macaques that intravenous inoculation with higher amounts of virus can lead to a more severe hepatitis (Tsarev et al., 1994b). Alternatively, the strain of virus may have been responsible; the virus used has been described as atypical (Chauhan et al., 1994) due to its ability to survive several cycles of freeze-thawing and exposure to 37°C for two days and because of its lack of reactivity with chimpanzee anti-HEV serum.

1.3.12   Infection in pregnancy

Studies on the severity and frequency of viral hepatitis during pregnancy, overall and in individual trimesters, have yielded conflicting results and controversy exists over whether pregnancy is a risk factor for development of clinically apparent viral hepatitis and a severe outcome of infection. Prior to the development of serological assays for HEV infection, the proportion of patients with NANB hepatitis due to HEV could not be assessed. In addition, the data for developing countries may have been affected by inequalities in the availability of healthcare and by the selective hospitalisation of only the most severe cases. There is now evidence that an important characteristic of HEV infection is the high case fatality (5-25%) and incidence of severe hepatitis in late pregnancy (Khuroo et al., 1981; Tandon et al., 1982; Khuroo et al., 1983; Kane et al., 1984; Myint et al., 1985; Zhuang et al., 1991; Tsega et al., 1992). Fulminant hepatitis can occur at any gestation but is most commonly observed in the later stages of pregnancy. In contrast, a recent review from India showed that

pregnant women who contract viral hepatitis did constitute a risk group for development of acute liver failure, however the mortality of this group was similar to that for non-pregnant women and men (Acharya *et al.*, 2000). Any increased risk of severe disease seen for HEV in late pregnancy cannot easily be compared to that for HAV and HBV, since outbreaks of the former are rare amongst adults in endemic areas, and outbreaks of HBV are limited by the mode of transmission.

Some studies have shown an excess of clinically apparent cases in women compared to men (Nepal, Kane *et al.*, 1984), however others have not (Myint *et al.*, 1985; Jaiswal *et al.*, 2001) or have shown the reverse (Naik *et al.*, 1992). This excess morbidity in non-pregnant women is not generally reported for other forms of viral hepatitis and the role of factors such as the degree of exposure and nutritional state, have not been defined. A higher attack rate has been suggested for HEV/ NANBH amongst pregnant women versus non-pregnant women of child bearing age (Khuroo *et al.*, 1981; Nayak *et al.,* 1989; Rab *et al.*, 1997), however, pregnancy may make a woman more likely to present to medical services during an outbreak of hepatitis. It should be noted that acute liver failure due to hepatitis E does occur in male adults and non-pregnant women and children. The differences observed between studies may represent true geographic variations of HEV strains or the presence of confounding factors in the relevant populations.

There may be underlying immunodeficiency in pregnancy resulting either from decreased IgG levels from haemodilution (Benster and Wood, 1970), decreased lymphocyte responses (Purtilo *et al.*, 1972; Kasakura 1971) and decreased lymphocyte numbers (Sridama *et al.*, 1982). These findings may explain the increased occurrence of herpes simplex virus hepatitis in the second and third trimesters of pregnancy (Kaufman *et al.*, 1997). Lower IgG levels have been observed in pregnant women with NANBH, when compared to non-pregnant women with NANBH and to pregnant women with HBV infection (Nayak *et al.*, 1989). In areas of poor nutrition, pregnancy may result in a relative malnourishment with an increased risk of severe hepatitis.

The mechanism of hepatocyte destruction in HEV acute liver failure is incompletely understood. The observation of atypical clinical features such as itching and joint pain may suggest an immune complex phenomenon. Given that the hepatocyte destruction following HEV infection may be immune mediated, it is unclear how a relative immune suppression could lead to a more marked hepatitis. The high mortality of

HEV infection in the later stages of pregnancy remains unexplained, but limited access to appropriate medical facilities in endemic areas may be contributory.

1.3.13   Foetal complications of maternal hepatitis E

There are few data on the vertical transmission of HEV. However, one study involving eight babies born to mothers infected in the third trimester showed HEV to cause intrauterine infection with substantial perinatal morbidity and mortality (Khuroo *et al.*, 1995). One baby had icteric hepatitis, four had anicteric hepatitis; two of these died, one of these from hepatic necrosis. The contribution of poor maternal health to these deaths is uncertain. HEV RNA was detected in either cord or birth blood in five babies and persisted for one month in two of these. In total, six babies showed evidence of infection, one by persistence of specific IgG past six months of age. The failure to detect IgM in three of these children may be due to non-production or lack of assay sensitivity, however, IgM was detected in the other three.

The effects of hepatitis E infection in early pregnancy are uncertain, however any serious infection in pregnancy may lead to an increased foetal loss or premature delivery. In one study of acute viral hepatitis in pregnancy, premature delivery was the commonest foetal complication observed (Medhat *et al.*, 1993).

1.3.14   Pathogenesis of hepatitis E infection

The pathogenesis and replicative strategy of HEV remains incompletely understood. Primate models of infection, and rare human volunteer studies, have been key in determining the current knowledge on the natural history, diagnosis and histology of hepatitis E. Transmission from human to cynomolgus macaque (*Macaca fascicularis*) was first described in 1983 (Balayan *et al.*, 1983). Other primates infected include green monkeys (*Cercopithecus aethiops*) (Anzhaparidze *et al.*, 1986), tamarins (*Sanguinus mystax*) (Inoue *et al.*, 1986), rhesus monkeys (*Macaca mulatta*) (Panda *et al.*, 1989), owl monkeys (*Aotus trivirgatus*) (Ticehurst *et al.*, 1992b), chimpanzees (*Pan troglodytes*) (Arankalle *et al.*, 1988), Pigtail macaques (*Macaca nemestrina*) and Squirrel monkeys (*Saimiri sciureus*) (Tsarev *et al.*, 1993b). Non-human primates have been infected by human HEV isolates from a wide range of geographical areas.

There has been a report of pigs (*Sus scrofa domestica*) developing hepatitis and jaundice following experimental inoculation with human hepatitis E (Balayan *et al.*, 1990). In contrast, pigs were not infected when inoculated intravenously with around ten monkey infectious doses of human HEV (Mexican and Pakistani strains) (Meng *et al.*, 1998a). Experimental inoculation of pigs with pig HEV did not induce illness or biochemical evidence of hepatitis (Meng *et al.*, 1998), despite evidence of replication and secondary transmission. A later study demonstrated that swine (US isolate) and human (US-2 isolate) HEV could infect pigs with evidence of viral shedding and histological evidence of hepatitis, although in the absence of clinical disease or biochemical markers of hepatitis (Halbur *et al.*, 2001). HEV has also been shown to infect rats (Maneerat *et al.*, 1996) and sheep (Usmanov *et al.*, 1994). Efficient replication of HEV in cell culture has only recently been described (see section 1.3.4).

1.3.14.1   Non-human primate models

The pattern of HEV infection may vary considerably depending on the species of non-human primate used. In addition, animals of the same species may respond differently within the same experiment and this makes generalisations regarding pathogenesis difficult.

Cynomolgus macaques are most reliably infected and develop the most marked hepatitis, therefore much of the data comes from studies involving their use (Longer *et al.*, 1993; Krawczynski *et al.*, 1989; Ticehurst *et al.*, 1992b; Tsarev *et al.*, 1993a; Van Cuyck-Gandre *et al.*, 1998). Furthermore, serial passage of virus through cynomolgus macaques can lead to increased severity of disease (Bradley *et al.*, 1987). The general pattern following inoculation is virus excretion in stools, viraemia and then hepatitis with seroconversion. Intravenous inoculation of the cynomolgus macaque results in an incubation period of around 2-5 weeks, with a peak ALT level around one week later. The acute phase of the illness lasts up to one week after the peak ALT.

One study involving several cynomolgus macaques (Longer *et al.*, 1993) extensively documented the pathology of HEV infection. Eight animals were infected, each exhibiting a variable onset and severity of disease. The general sequence of events in these animals appeared to be a rise in ALT at 10-19 days post-inoculation and around that time histological changes appeared in the liver and HEV antigen was seen in the

hepatocyte cytoplasm, and virus particles were detected in bile. Pathological changes were consistent with acute viral hepatitis and included diffuse swelling of hepatocytes with focal cytoplasmic bile stasis and occasional hepatocyte necrosis. Mild, predominantly lymphocytic infiltrates were noted. Four animals were sacrificed at eight days after the initial rise in ALT. Serum anti-HEV was initially detected (days 27-39) in the remainder before or coinciding with the peak in ALT (days 28-39). On day 39 hepatic HEV antigen could no longer be detected but pathological changes persisted. These findings agreed with previous work by other investigators.

Other investigators have documented viraemia in addition to HEV particles in the bile and stool and detectable hepatic HEV antigen in the second or third week after inoculation, followed by the onset of the ALT elevation and pathological changes (Tsarev et al., 1993a). The outcome of experimental infection can be linked to the dose of virus given (Tsarev et al., 1994b), and this may have been responsible for some of the variation between studies.

The concomitant presence of HEV antigen and HEV RNA in liver and bile indicate viral replication and release of virions into the biliary system. HEV antigen appears to be located randomly in the cytoplasm of cells and without zonal preference in the liver lobule (although this relates to intravenous infection). Further evidence for HEV replication in the liver of other experimentally infected animals (rhesus monkeys) was provided later by the detection of HEV specific negative strand RNA in that tissue (Nanda et al., 1994). A recent study involving experimentally infected rhesus monkeys, demonstrated the relationship between the presence and location of HEV RNA in the liver and the detection of VLPs in bile, biochemical evidence of hepatitis, and the development of HEV antibody (Kawai et al., 1999). Animals were studied during the incubation, hepatitis and recovery period. During the incubation period, HEV RNA was detected in the cytoplasm of hepatocytes, particularly in the submembranous region, although the liver showed no histological abnormality and there were no VLPs in bile. One animal killed following detection of VLPs in bile, with normal ALT, had HEV RNA located only near the cell membrane of bile epithelial cells on the canalicular side. HEV RNA could not be detected in the liver of another animal sacrificed at the same stage of infection. During acute hepatitis, focal necrosis was observed in the liver but there was no HEV RNA detected in or around those areas. The bile epithelial cells were positive for HEV RNA, located throughout the cytoplasm.

The histology of infected liver tissue from different studies of experimental HEV infection, often shows hepatocyte necrosis, inflammatory cell and lymphocyte infiltrates and Kupffer cell hypertrophy. These changes are consistent with, but not specific to, HEV hepatitis.

There are conflicting data on the role of antibody in the disease process and the findings may not reflect the effects of natural infection, since animal data comes from intravenous challenge. Such intravenous inoculation can lead to the presence of HEV antigen in hepatocytes within seven days. Antibody was often detected first around the time of the most marked hepatitis (Tsarev et al., 1992; Ticehurst et al., 1992b; Longer et al., 1993; Kawai et al., 1999). The possibility that antibody was present earlier, but at low level, cannot be excluded. This was the case in four cynomolgus macaques infected with an HEV from Chad, when specific IgM was detected within 17 to 24 days following intravenous inoculation, a few days prior to the rise in ALT (Van Guyck-Gandre et al., 1998). Patterns linking antibody production to liver damage and viral clearance have been observed (Tsarev et al., 1993a and 1993b). In contradiction, some animals have shown an uncoupling of antibody production and liver injury or a biphasic illness (Bradley et al., 1987; Tsarev et al., 1993a). The study by Tsarev in 1993 found one macaque out of the five had a small rise in ALT around days 10-14 post-inoculation associated with virus in faeces. The second peak of ALT occurred at day 55 and was associated with viraemia and seroconversion and a marked hepatitis. Since HEV RNA was detected in bile/faeces of experimentally infected macaques at day 6 post-inoculation and in serum at day 9, preceding the initial rise in ALT, the biphasic pattern may have represented early viral replication rather than the possibility of another hepatotropic virus in the inoculum.

The delayed appearance of anti-HEV until after pathological changes in the liver noted in some experiments suggests antibody was not responsible for initiating these changes, or that the tests were insensitive. Such antibody may be linked to the more marked pathology seen as virus is eliminated. The suggestion that hepatocyte damage is an immune-mediated phenomenon was supported by the demonstration of cytotoxic/suppressor phenotype lymphocytes in the hepatic inflammatory infiltrate of HEV infected cynomolgus macaques (Soe et al., 1989). However, Balayan (Balayan et al., 1991) found that immunosuppression of macaques using steroids or cyclosporine did not alter the severity of the hepatitis, but did lead to prolonged virus excretion. Furthermore, the detection of HEV in faeces prior to illness indicates that

replication occurs in the absence of hepatocyte damage. These events are similar for virus from different geographic regions.

Anti-HEV antibody has been found in pre-inoculation sera from macaques and rhesus monkeys (Tsarev *et al.*, 1995) and in wild-caught Old World monkeys (Arankalle *et al.*, 1994b), possibly pointing towards previous exposure to HEV or a similar agent. The presence of low-level undetectable antibody or immunity may also account for variability between studies and the much more brisk HEV antibody response that has been seen in some animals following experimental infection.

The range of responses to experimental infection seen in non-human primates may reflect the varying outcomes of human infection (subclinical, mild to moderate illness, or acute liver failure). Despite the detailed analysis of experimental hepatitis E in non-human primates, several factors limit their application to the human form of the disease. These include lack of infectivity by the oral route, a mild or absent hepatitis (fulminant hepatitis is not seen) and lack of clinical signs equivalent to those in humans, and the absence of a cholestatic histology sometimes seen in human disease. Interestingly, a high titre oral dose of HEV, which caused relatively severe infection in cynomolgous macaques by the intravenous route, resulted in no apparent disease (Tsarev *et al.*, 1994b). Hepatitis E in pregnant rhesus monkeys was no more severe than in control animals (Tsarev *et al.*, 1995), however there was some evidence that pregnancy may have reduced the incubation period independent of the viral dose (Arankalle *et al.*, 1993c).

1.3.14.2   Pathogenesis of human infection

The pathogenesis of human HEV infection is not fully understood. The role of virus replication and the immune response in causing hepatitis has been debated. Several studies have shown variation in the dynamics of virus excretion. Virus has been detected by IEM in stool nine days prior to the onset of jaundice (Ticehurst *et al.*, 1992a). In several other stools from the same study, virus was no longer present when tested between days 8 and 15 post-icterus. Faecal viral concentrations were low; only one sample attained five particles per EM grid square (equivalent to 10^8 HEV/g). This low level shedding prior to and into the first week of jaundice contrasts with other viruses such as HAV, which is excreted in high amounts mainly before jaundice

occurs, and Norwalk virus, which may exceed 100 VLPs/ grid sq. in the early period after onset of the illness. In general, HEV has been detected in stool (using IEM and RT-PCR) from around one week prior to, and up to two weeks after, clinical symptoms (Balayan *et al.*, 1983; Ticehurst *et al.*, 1992a; Chauhan *et al.*, 1993; Chobe *et al.*, 1997; Zanetti *et al.*, 1999; Aggarwal *et al.*, 2000). In a prospective study focusing on 20 patients with acute hepatitis E, viraemia and faecal excretion of virus was not detected in 19 after biochemical resolution of hepatitis, and only one patient had viraemia alone after that time. However, there is now good evidence for protracted faecal shedding (suggested by the presence of HEV RNA) of virus of up to 52 days (Nanda *et al.*, 1995) and a prolonged viraemia of up to 112 days (Nanda *et al.*, 1995; Clayson *et al.*, 1995). These occurred despite clinical and biochemical recovery, and in some cases after a prior negative result, raising questions about the mechanism of viral persistence, release and tissue tropism. Viral release from hepatocytes without cell lysis or the presence of virus in alternative tissues could explain these observations. The role of immunodeficiency remains unknown in these cases.

If common, prolonged viral excretion after clinical recovery may be important in maintaining transmission during outbreaks. However, given the sensitivity and the target for RT-PCR, the presence of HEV nucleic acid in stool does not necessarily indicate infectivity.

HEV viraemia is present before the onset of symptoms (Chauhan *et al.*, 1993) and persists into the acute phase of the illness, usually for 1 to 2 weeks (Clayson *et al.*, 1995b; Zanetti *et al.*, 1999). The detection of HEV RNA in serum indicates either continued viral replication or possibly the release of genetic material from lysed hepatocytes. The presence of HEV RNA in stool generally parallels the viraemia.

The pathogenesis of infection is complicated by these and other unusual features found in some cases. Some patients appear to develop jaundice with HEV viraemia yet have no antibody response (Clayson *et al.*, 1995b). HEV causing hepatitis in a Nepalese patient without an antibody response also produced antibody negative infection in cynomolgus macaques, pigs, and rats (Clayson *et al.*, 1995b). The negative antibody result could be a function of the assay design, or be due to circulation of an atypical strain of HEV, however it may help explain the unusual epidemiological features of HEV infection (see section 1.3.5).

Liver tissue is rarely available from HEV infected patients with uncomplicated disease; samples which have been examined showed both cholestatic and classic acute

viral hepatitis (Gupta and Smetana, 1957; Khuroo, 1980). The histological picture of human disease is different to that of the non-human primates (which show inter-species variation)- neither cholestasis nor fatty transformation has been observed in non-human primates. The intralobular infiltrate was predominately lymphocytic but polymorphonuclear leucocytes were seen also. The standard picture was of focal hepatic necrosis, ballooned hepatocytes, acidophilic degeneration of hepatocytes with an inflammatory infiltrate dominated by cytotoxic lymphocytes. Some examples of submassive or massive hepatic necrosis have been recorded, particularly in association with fatal infections. In two cases of fulminant hepatitis E, viral antigens were detected only in the cytoplasm of hepatocytes without any significant inflammatory infiltrate (Lau et al., 1995). Since the hepatocytes were degenerating, it was suggested that the virus was cytopathic, however these patients are not representative of the general population since one was a child with another liver disease (Wilson's disease) and the other was 20 weeks pregnant.

1.3.15   Co-infections and super-infections of HEV with other hepatitis viruses

Since HEV is endemic in areas where other hepatitis virus infections are common, co-infections and super-infections between these agents are likely to occur readily. In countries such as India where almost all adults are immune to hepatitis A, HEV and HAV co-infections are most likely in children (Malathi et al., 1998). Adults are more likely to acquire HEV infection in the presence of chronic infection with HBV (Khuroo et al., 1994) or HCV (Nanda et al., 1994). Clearly, HEV infections complicated by underlying liver disease or acute hepatitis A or B may not follow a typical course. One study has suggested that children co-infected by HAV and HEV are more likely to develop liver failure than those with single infections (Arora et al., 1996). However, in those cases, detection of HAV IgM was used to indicate acute HAV infection and since this may remain positive for several months after recovery, an actual co-infection could not be certain. Whilst co-infection with different hepatitis agents may also lead to more severe disease in adults (HAV and HEV, Zanetti et al., 1999), this is not always the case (HEV and HBV, Khuroo et al., 1994). In comparison, hepatitis A superinfection of chronic hepatitis C, can lead to a more

severe illness than that seen in individuals without existing liver disease or who are hepatitis B carriers (Vento *et al.*, 1998).

1.3.16   Protective immunity

The correlates of protective immunity against HEV have not been clearly defined. Similarly, the natural history of protective immunity against hepatitis E infection is uncertain, given the continued occurrence of outbreaks in endemic regions. Persisting HEV IgG in adults has been proposed to confer resistance to re-infection (Bryan *et al.*, 1994) in the natural setting, whereas human antibody resulting from distant HEV infection did not protect monkeys from hepatitis after intravenous challenge with the isolate from the original human infection (Chauhan *et al.*, 1998). However, in this study the onset of hepatitis was delayed, with one of the three rhesus monkeys developing a raised ALT/AST on day 86 post-challenge. The incubation period is usually around 9-12 days in monkeys following intravenous challenge. The implication of this long incubation period is unclear since control animals also had a prolonged incubation.

Experimental HEV infection of animal models indicates that protection remains against heterologous HEV challenge. Rhesus monkeys infected with an HEV from the former USSR were assumed to be immune to later challenge with an Indian strain, however, this was on the basis of no rise in ALT alone (Arankalle *et al.*, 1993b). A later study, using rhesus monkeys previously infected with an Indian HEV, showed absence of viral replication using RT-PCR of bile and stool following intravenous inoculation with different HEV isolates from India (Arankalle *et al.*, 1995b). None of the ten monkeys showed an increase in HEV IgG following exposure to HEV, however, two did show a short-lived increase in ALT at 30 and 37 days post-inoculation. The authors assumed this transient evidence of liver damage to be unrelated to HEV in light of the RT-PCR findings. The animals used in this study all had relatively high HEV IgG titres ($\geq$1:1,000), and this may have been critical to the outcome, since wild-caught rhesus monkeys with detectable HEV IgG have shown an antibody titre dependent response to experimental HEV challenge (Arankalle *et al.*, 1994b). Two of these wild-caught monkeys, one with an HEV IgG titre of 1:1,000 and one with a titre of 1:10,000, appeared to be resistant to infection. Of two monkeys with pre-challenge titres of 1:100, one developed a rise in HEV IgG titre after

challenge (with no evidence of hepatitis or virus excretion in stool) and more importantly, one developed a typical HEV infection. This possible antibody titre dependent protection was evaluated in a later study using rhesus monkeys infected with HEV five years previously (Arankalle *et al.*, 1999). One monkey with an HEV IgG titre of 1:50 was protected against infection and disease but another animal with an HEV IgG titre of 1:400 showed evidence of infection (stools HEV RNA positive) without disease. Therefore, antibody titre may play a role in protection against future exposure to HEV in monkey models and this may also be the case in human HEV infection. It should be noted that a decline in antibody titre to undetectable levels following natural HEV infection would not necessarily result in susceptibility, since cellular immunity may have been effectively primed.

Repeated HEV infection in humans has not been described, although clearly this does not exclude such an event. The ability of human HEV IgG antibody to neutralise HEVs from different countries was investigated using a cell culture based assay (Meng *et al.*, 1998). This assay used hepatocarcinoma cells (PLC/PRF/5) which support HEV RNA production without CPE. HEV antibody from individuals infected across parts of the world, neutralised the infectivity of HEV strains from Burma, Pakistan, Mexico and Morocco using this system. However, inconsistencies were observed when serum samples of Asian origin were used against Mexican or Moroccan HEV. Although this could have been a dose effect, it does raise the possibility of antigenic differences as a factor. Whether the immunity stimulated by prior infection can be reproduced by immunisation using only the putative structural proteins of HEV as antigens remains to be determined.

1.3.17   Passive immunisation

The presence of existing naturally acquired HEV IgG may offer protection against disease in the epidemic setting (Bryan *et al.*, 1994), therefore passive immunisation might be expected to be protective. The use of immune globulin prophylaxis for hepatitis E appears ineffective in humans (Joshi *et al.*, 1985; Zhuang *et al.*, 1991a; Khuroo and Dar, 1992), although this may have been due to the doses used. One study of immune globulin prophylaxis of pregnant women suggested a beneficial effect but this did not achieve statistical significance (Arankalle *et al.*, 1998). Passive immunoprophylaxis may modify infection in animal models (Tsarev *et al.*, 1994a)

although this was not confirmed (Chauhan *et al.*, 1998). In the study by Tsarev (Tsarev *et al.*, 1994a), two macaques had passively acquired HEV antibody titres of 1:40 and two had titres of 1:200 at the time of intravenous HEV challenge. Following inoculation, all four animals had viraemia and HEV in stool (by RT-PCR) and one of the animals with an HEV antibody titre of 1:40 developed hepatitis. In order to achieve an antibody titre of 1:200 by passive transfer, Tsarev and colleagues had to transfuse animals with around 10% of their blood volume using high titre HEV convalescent serum. This protocol, or one of similar efficacy, is unlikely to be transferable to humans.

1.3.18 Active immunisation

Since endemic areas may not have the facilities for large scale production of HEV immunoglobulin, a vaccine would be more appropriate. Currently there is no available vaccine for routine human use. Preliminary studies have focused upon immunogens derived from ORF 2, expressed in bacteria (Purdy *et al.*, 1993; Li *et al.*, 2000) and eukaryotic cells (Tsarev *et al.*, 1994a) or presented as DNA immunisation (Panda *et al.*, 1997, He *et al.*, 1997). Immunisation of rats with an ORF 2 subunit vaccine (C-terminal 267 amino acids) has generated an antibody response with similar specificity to that found after natural infection (Li *et al.*, 2000). In addition, immunisation of mice using DNA coding for sections of putative HEV protein has been shown to induce immunological memory (He *et al.*, 2001). Vaccines based upon self-assembled capsid proteins expressed in insect cells may be utilised in the future, as this approach has been successful for rabbit haemorrhagic disease virus (Laurent *et al.*, 1994). The efficacy of candidate vaccines varies according to the criteria used to indicate successful immunisation. Several studies have shown vaccine induced protection against hepatitis in monkeys following intravenous challenge (Purdy *et al.*, 1993; Tsarev *et al.*, 1994a; Tsarev *et al.*, 1997) and this has included heterologous HEV challenge (Yarbough *et al.*, 1997; Tsarev *et al.*, 1997). However, there is some evidence from the cell culture based assay described earlier, that protection against all strains may be difficult to obtain, since antibodies generated by immunisation of rabbits using the Burmese strain ORF 2 product cannot neutralise the Mexico or Morocco strains of HEV (Meng *et al.*, 1998). This effect may have been the result of antigenic heterogeneity between strains. Alternatively, the effect may have resulted

from antigenic masking of important epitopes due to the protein structure, as the same study showed that antibody from guinea pigs immunised with a fragment of the ORF 2 protein (C2, amino acids 225-660) could neutralise the Mexican strain but not the Moroccan. However, serum taken from cynomolgus macaques experimentally infected with either the Burmese, Pakistani or Mexican strain of HEV could neutralise each of those strains and one from Morocco. This provides some evidence that immunity generated by HEV infection is qualitatively different from that stimulated by immunogens derived from ORF 2.

In many immunisation experiments, hepatitis E RNA or virus has been detected in the stools from a proportion of animals following subsequent challenge, suggesting viral replication had occurred (Purdy et al., 1993; Tsarev et al., 1994a; Tsarev et al., 1997). Immunisation of monkeys with a trpE fusion protein containing the carboxyl two-thirds of the putative capsid protein from the Burma strain (Purdy et al., 1993) was protective against hepatitis after intravenous challenge with Burmese and Mexican strain HEV. However, the animal challenged with the Mexican strain was found to have HEV RNA in stool and HEV antigen in liver tissue. These findings are important, since viral replication in the absence of hepatitis has been observed in unvaccinated monkeys following experimental challenge (Jameel et al., 1993).

Although hepatitis E has a relatively long incubation period, post-exposure immunisation at 48 hours using a recombinant, baculovirus expressed, ORF2 (Pakistani strain) based vaccine (Tsarev et al., 1997) was unsuccessful. Although disease was not prevented, there may have been a reduction in duration of viraemia and virus excretion.

It should be noted that all immunisation studies have relied upon intravenous challenge with HEV to demonstrate protection against infection, and it is unclear how this affects their relevance to the natural situation of infection by the oral route.

Given that the HEV vaccines studied so far may not prevent viral replication in the experimental host, the use of any human HEV vaccine may not protect against viral excretion sufficiently to prevent epidemics occurring. Furthermore, although experimental infection of monkeys can confer complete protection against replication following challenge with different HEV strains (from the same geographical area) (Arankalle et al., 1995b), this is not necessarily the case for human infection. Under these circumstances, initial immunisation programs might be best targeted to preventing disease in high risk groups, such as women of child-bearing age. Despite

the potential difficulties associated with successful immunisation against HEV, human vaccine trials are underway in Nepal (Stevenson, 2000).

In the absence of an effective vaccine, the incidence of hepatitis E could be reduced by ensuring that water purification measures are adequate and reliable.

1.4   Acute liver failure

Acute liver failure (ALF) (reviewed in Caraceni and Van Thiel, 1995) is a clinical syndrome comprising altered consciousness (encephalopathy) and impairment of the detoxifying and synthetic functions of the liver, resulting from massive necrosis of hepatocytes. Even with intensive medical therapy, patients with marked encephalopathy have a poor prognosis with only 10-40% surviving. The term fulminant hepatic failure (FHF) was originally defined by Trey and Davidson in 1970 as the consequence of severe liver injury leading to encephalopathy within eight weeks of the appearance of first symptoms, in the absence of pre-existing liver disease. It is now generally accepted that FHF can occur in patients with some degree of underlying liver impairment e.g. Wilson's disease. Further modifications or additions to the early description were provided by Bernuau (Bernuau et al., 1986), where FHF is used to describe patients with encephalopathy within two weeks of the appearance of jaundice; when encephalopathy results 2-12 weeks after jaundice this was classified as subfulminant hepatitis. Gimson defined late-onset hepatic failure as encephalopathy occurring 8-24 weeks after the appearance of jaundice (Gimson et al., 1986).

In the UK, the commonest cause of FHF is paracetamol overdose, but in many other countries the hepatitis viruses are most commonly implicated. Less commonly associated causes in the UK include infection with herpes simplex viruses 1 and 2, varicella-zoster virus, Epstein-Barr virus, and cytomegalovirus. However, even after excluding hepatotoxins (including therapeutic drugs), no specific aetiology is found in a significant proportion of FHF cases.

Although the mechanisms of viral FHF are not fully characterised, it is thought that an excessive immune response may play a major role.

1.5 The concept of molecular pathogenesis

Clearly the cellular tropism, virulence and pathogenesis of viruses are dependent upon their genetic composition. However, the outcome of infection is also influenced by host factors. Viruses may undergo genetic change as a result of selective pressure on random mutations introduced during replication. Such pressure may result from host immunity or the use of antiviral drugs or immunisation. In addition, those changes in the viral genome which have little or no effect on interaction with the host may accumulate over time; typically these are silent mutations, where a change in the third base of a codon does not alter the amino acid produced. There are often several genome variants of a single strain or species of virus circulating at one time, both in the population, and within an individual for some viruses. This is particularly true for viruses with RNA genomes, since RNA dependent RNA polymerases and reverse transcriptases lack error correction domains (proofreading function). Such genome variation may lead to an altered pathogenesis, and it is this possibility that forms the focus of this thesis. Could the severity of HEV infection be linked to a viral factor rather than the host response?

In the absence of specific data it is reasonable to assume that the virulence or pathogenicity of a particular viral strain could be linked to replicative efficiency. Studies may therefore focus upon putative transcription and translation control regions of viral genomes. There are examples of how changes in such regions can influence the pattern of resultant disease following infection. Analysis of the Sabin vaccine strain of poliovirus indicated neurovirulence to be linked to nucleotide changes in the 5' UTR of the genome, an area which functions as an internal ribosome entry site (Minor, 1992). Similarly, the cardiovirulence for mice of Coxsackie B3 has been mapped to nucleotide changes in the 5' UTR (Gauntt and Pallansch, 1996).

Another target for altered virulence may be the specificity of cell receptor binding proteins. Once again the enteroviruses provide an example of how an amino acid change in a capsid component may affect virulence; changes in the VP1 of Coxscakie B4 can influence the virulence phenotype (Caggana et al., 1993).

Such factors are likely to contribute to the clinical differences following infection with serologically identical strains; outbreaks of Enterovirus 71 have been associated

with a relatively benign hand, foot and mouth syndrome or with meningitis and paralysis (Zheng *et al.*, 1995).

Host immunity also affects the outcome of infection; immunodeficient states affect both the acquisition and course of infections. Additionally, it is well recognised that infection in adulthood by a virus that usually first infects in childhood can lead to more severe disease. Examples include hepatitis A and varicella-zoster virus. A more vigorous immune response could result in a more marked disease where the mode of injury is immunopathological, such as hepatitis A. A more specific example of a host factor mediating resistance to infection can be found in the HIV 1 receptor gene. The 32 base pair deletion mutant of the human CCR5 allele confers partial protection against HIV 1 infection (Huang *et al.*, 1997). More subtle, uncharacterised differences between the immune function and genetic composition in immunocompetent individuals may also be important.

1.6 AIM

Hepatitis E virus usually causes a relatively mild, self-limiting hepatitis. Occasionally, infection may progress to acute liver failure. The outcome of infection is determined by the interplay between both viral and host factors. The aim of the project was to attempt to identify any viral factors that may be linked to the progression of infection to acute liver failure.

This was to be achieved by nucleotide sequencing of viral nucleic acid derived from infected liver taken from a case of acute liver failure, and comparison with published data from other isolates, targeting areas suspected to be of importance in regulation of viral replication. This project was proposed at a time when the published HEV sequences were derived from virus passaged in monkeys, and it was unclear whether this may have led to alteration in viral characteristics.

Subsequent expression *in vitro* of viral genes was to be used to provide further information on the proteins generated.

Overview of methods

**Total RNA extraction
from liver tissue**

**Northern blot
analysis**

**First strand cDNA
for standard PCR**

**First strand cDNA
for full length PCR
and RACE**

**cDNA for
plasmid
library**

**Homopolymer
tailing and RACE**

**Full length
HEV PCR**

**Full length
HEV clones**

**Subcloning
sequencing**

**Individual
ORF 2, 3
PCR**

**Cloning and
Sequencing**

**HEV consensus
sequence**

**Expression
studies**
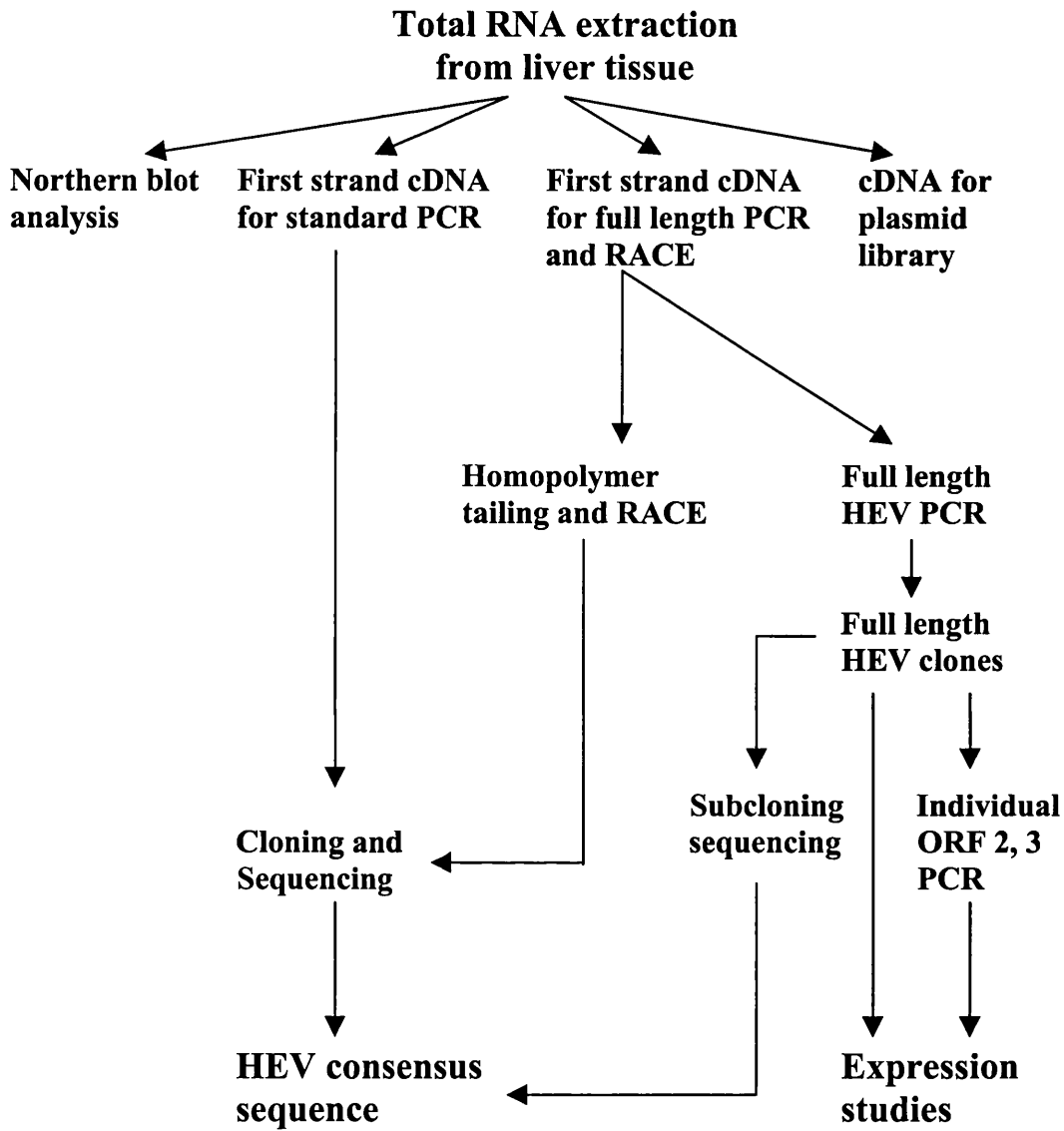
Figure 2.1: Diagram showing the basic experimental plan used to derive the study strain consensus sequence and for generation of clones for expression studies

CHAPTER TWO

MATERIALS AND METHODS

2.1   Source of liver tissue

The project was based upon a case of acute liver failure resulting from hepatitis E infection acquired in India. KS was a 57 year old man who travelled extensively in India over a period of 2 months. He became ill 3 days after returning to the UK, initially presenting to his general practitioner with non-specific symptoms including fever and anorexia. He was prescribed antibiotics. Eight days later he attended an accident and emergency department suffering from nausea and vomiting; at the time he was jaundiced. He was admitted to hospital with a provisional diagnosis of hepatitis A. Biochemical tests on admission revealed a biblirubin of 440μmol/l and an AST of 4479 IU/ml. Eight days after admission he developed hepatic encephalopathy and was transferred to another hospital where a liver transplant was performed. Death occurred 33 days post-operation following cardiac arrest. His explanted liver showed severe acute hepatitis with bridging necrosis; samples were stored at -70°C. Serological tests excluded both hepatitis A, B and C, and later testing for HCV RNA was negative. Antibody to hepatitis E (Fagan *et al.,* 1994) was detected and hepatitis E genomic material was detected in the explanted liver.

KS lived in Nairobi and Uganda prior to entry to the UK in 1972. His trip to India included visits to Bombay and Bangalore, accompanied by his wife, and trips to Goa and Delhi with an Indian friend. There was some local flooding during the time he spent in Bombay. His wife had an uncomplicated episode of diarrhoea 4 days prior to returning to the UK, but did not have or subsequently develop a hepatitis.

2.2 Total RNA extraction from human liver

This procedure was kindly performed by Dr E A Fagan and was based upon guanidinium isothiocyanate extraction and separation on a caesium chloride gradient (Sambrook *et al.,* 1989).

Around 500mg of liver tissue was processed. The RNA was quantified by measuring the optical density at 260/280nm and aliquoted into 5μg lots and stored under ethanol at -70°C.

2.3 First strand cDNA synthesis

Two methods were employed to generate cDNA for use in standard length amplicon PCR: random-primed synthesis and oligo-dT primed synthesis. Prior to use in PCR the first strand cDNA was purified as described in method 2.5.

The production of cDNA for use in the full length genomic PCR and RACE experiments is detailed in method 2.10.1.

2.3.1 Random primed synthesis

5μg of total RNA ethanol precipitate was spun in a microcentrifuge for 45 mins. The pellet was resuspended in 10μl DEPC TE pH 7.6 (appendix 10), heat denatured at 70°C for 1 min and placed on ice. The following were added to the RNA on ice: 1μl HPRNI (100 units, Amersham Intl, UK), 31μl first strand synthesis buffer (50mM TrisHCl pH8.3, 8mM $MgCl_2$, 50mM NaCl, 1mM DTT, 1mM each dNTP, 0.2μg random hexamers), 0.3μl ^{32}P dCTP (3000Ci/mmol), 5.7μl DEPC water and 2.0μl Super RT AMV reverse transcriptase (30 units/ul, Bioquote Ltd, UK).

The reaction was mixed, 1μl taken as baseline for later analysis and then incubated at 42°C for 2.5 hrs, after which a further 1μl was taken as end point for analysis. The reaction was stopped using 3μl 0.5M EDTA.

2.3.2 Oligo-dT primed synthesis

A 5μg pellet of total RNA was dissolved in 13.5μl of DEPC water and 10μl of 1mg/ml oligo-dT$_{12-18}$ was added. The mixture was heated to 70°C for 2 mins (to denature the RNA) and cooled on ice.

On ice: The following were mixed- 10μl 5x first strand buffer (supplied with enzyme), 0.5μl 0.1M DTT, 10μl dNTPs (5mM each), 1μl HPRNI (100 units, Amersham Intl, UK) and 5μl Murine RT (200u/μl, Life Technologies). This was added to the RNA:primer mix and incubated at 37°C for 1 hour.

A further experiment was performed using an HEV specific primer (table 2.1, set 2 primer, position 7174) substituted for oligo-dT.

2.4 Analysis of first strand cDNA reactions

The addition of radiolabelled dCTP to the cDNA synthesis reaction allowed the percentage incorporation of trichloroacetic acid precipitable counts to be determined. Total and precipitable counts were used from reaction baseline and end-point in the calculation.

Alternatively, a filter binding assay was used to determine the percentage incorporation of radiolabel. Reaction aliquots were adsorbed onto Whatman DE81 paper and dried. Washing the paper with 0.5M sodium phosphate removes unincorporated nucleotides.

2.5 Purification of cDNA

First strand cDNA was purified by two methods: solvent extraction with ethanol precipitation or using a silica based commercial kit.

2.5.1 Solvent extraction and ethanol precipitation

The cDNA reaction mix was extracted with an equal volume of TE (pH8.0) saturated phenol:chloroform (appendix 10), and the upper aqueous phase transferred into a fresh tube. Further extraction with an equal volume of chloroform:isoamyl alcohol (24:1)

was done and the upper aqueous phase saved. An equal volume of 4M ammonium acetate and twice the combined volumes of cold ethanol was added and the mixture chilled for 15 mins on dry ice, then warmed to room temperature with gentle shaking to dissolve any unincorporated dNTPs. This was microcentrifuged for 10 mins and the supernatant carefully removed. The pellet was washed once at room temperature using 50μl 2M ammonium acetate plus 100μl cold ethanol by applying gentle shaking, then centrifugation and aspiration. The pellet was washed with 200μl cold ethanol, spun for 2 mins in a microcentrifuge and the supernatant removed. The pellet was air dried and redissolved in 40μl water.

2.5.2 Purification of cDNA using silica particles (QiaexII kit)

The QiaexII (Diagen, GmbH) kit can be used for purification/desalting of DNA solutions and extraction of DNA from agarose or polyacrylamide gels. The kit uses the principle of adsorption of DNA to silica-gel in the presence of high salt. Agarose is solubilized by a chaotropic agent and heating to 50°C. Incubation of DNA in a highly electrolytic environment with large anions provided by chaotropic salt forces the DNA to bind to the silica particles. After ethanol washes the DNA can be eluted in low salt solutions/water. DNA fragments between 40 bp and 50 kb can be recovered. The kit was used according to manufacturer's protocols.

2.6   Standard length amplicon HEV PCRs

The conditions used for the successful PCRs to amplify sections of the HEV genome are given in table 2.2; a list of all primers is given in table 2.1. All reactions were set up under clean room conditions and analyzed in a separate laboratory.

Some primer sets were kindly supplied by S K Panda, having previously been manufactured in India and validated on HEV isolates from India.

Primer set	3' HEV position[1] / polarity/ length	Sequence (5'-3')
SET A	4858 /-/ 18	CAGG G/C CCCCA A/G TTCTTCT
	4477 /+/ 20	GC T/C ATtATGGA G/A GAGTGtGG
	4836 /-/ 28	gcgaagcttTaAGCCGGCCGGCGAAGCG
	4538 /+/ 26	gcgctgCaGCGTGGATCtTGCAGGCC
SET 1	6282 /+/ 20	cgggatccGCTGACACTCTG
	6749 /-/ 20	GAGCGGCAGGACaAAGAAGG
	6724 /-/ 20	GcgaAtTcCTGGATGGTGGA
SET 2	6653 /+/ 20	gCgGattCCaTGGTTAATGT
	7174 /-/ 20	CgGaattCGCGgAACGCAGA
	6704 /+/ 19	cgGAtccGgaCcAGGTCAC
SET 4	4798 /+/ 22	cgGaatTcTATGCAGGTGTTGT
	5311 /-/ 22	CgAAGGGGTTGGTTGGATGAAT
	5282 /-/ 22	gcgGatccGGGCTGAGAATCAA
SET 8	5111 /+/ 21	tTGTGTCGGGTGGAATGAATA
	5608 /-/ 21	aCTAAGAGGGGCGGCATAAaG
	5145 /+/ 23	GCgGatCCCATGGGTTCGCGACC
	5571 /-/ 27	GcGCtGcaGGCCACGGaAGAGGTaAGg
SET 10	6068 /+/ 24	GGAcTTTGCCCTTGAgcTTGAGTT
	6331 /-/ 24	TGAGACaACGGGACGGGAGTAGAA
SET 11	25 /+/ 26	aGGCAGACCACaTATGTGGTCGATGC
	456 /-/ 22	AAAACCCGTCGAgGCAGTAAGT
	516 /-/ 26	GATGGTGACATATCATGaAGGGAgTA
SET 13	2176 /-/ 24	GCggAtccatAAAACCTAAGTCAG
	1437 /+/ 24	TcCGTAAGGCGcTCTCAAAGTTTT

Table 2.1: Primers used in successful short amplicon HEV PCRs

[1]Numbering is based on the consensus nucleotide sequence for the studied HEV isolate. Lower case denotes a mismatch with the consensus sequence; often resulting from creation of restriction endonuclease sites to aid cloning.
Set A- reference Ray *et al.*, 1991. Set 8- amplicon contained a deletion, see results section 3.1

Primer set (Table 2.1)	Taq polymerase	Target cDNA /µl	Reaction volume /µl	Cycling parameters (Techne PHC3 cycler)
set 1 semi-nested	1.5u 1.5u	2.0-5.0 (a) na	50 50	95°C 1min/53°C 1min/72°C 1.5 min x30 cycles for both
set 2 semi-nested	1.5u 2.0u	2.0 (a) na	50 100	95°C 1min/53°C 1min/72°C 1min x30 cycles for both
set 4 semi-nested	1.5u 2.0u	2.0 (a) na	50 100	95°C 1min/53°C 1min/72°C 1min x30 cycles for both
set 8 (d) nested	1.5u 2.0u	3.0 (b)	50 100	94°C 1.2min/55°C1.2min/72°C 1.5min x30 cycles first round, 35 cycles nested
set 10	2.5u	2.0 (a)	100	94°C 1min/52°C 1.2min/72°C 1.5min x40 cycles
set 11 semi-nested	1.5u 2.5u	2.0(a) or (b) na	50 100	94°C 1min/52°C 1.2min/72°C 1.5min x40 cycles for both
set 13 (c)	1.5u	2.0 (a)	50	94°C 1min/52°C 1.2min/72°C 1.5min x40 cycles for both
set A nested	2.5u 2.5u	2.0 (a)	50 50	94°C 1min/ 50°C 2min/ 72°C 1.5min x30 cycles for both

Table 2.2: Reaction parameters of short amplicon HEV PCRs with clonable products

(a)- cDNA primed by random hexamers.
(b)- cDNA primed by oligo-dT.
(c)- semi-nested reaction failed.
(d)- amplicon contained an internal deletion.

All reactions contained 200µM dNTPs, 0.4µM each primer, 1.5mM magnesium chloride, and buffer comprising 50mM KCl, 10mM TrisCl pH 8.3. Each round terminated in a 10 min 72°C extension step. Taq polymerase supplied by AB Ltd, UK.
Typically, one tenth of the first round reaction was used as target in the second PCR.

2.7 Cloning of standard length HEV PCR products

Standard cloning techniques were used in addition to proprietary cloning kits. Many of the PCR primers were designed to include restriction endonuclease sites to aid cloning. Vectors employed were pBluescript II and pCR Script phagemids (Stratagene Inc, USA.) and pCRII/2.1 TA cloning phagemids (Invitrogen BV, The Netherlands). Details of the plasmids are given in appendices 2-4. All the cloning vectors used allowed blue/white selection of potential recombinant clones. Insertion of DNA into the multiple cloning site within such plasmids disrupts the expression of β-galactosidase by lacZ thereby preventing the production of blue colour from X-gal substrate under IPTG induction.

2.7.1 Cloning into pBluescript (Stratagene Inc., USA)

The pBluescript phagemids were designed to simplify commonly used cloning and sequencing procedures, allowing production of single-stranded DNA and possessing an extensive polylinker available in two orientations. Ligation reactions were done according to manufacturers guidelines (also see appendix 9). Transformation of bacteria was done according to the method set out in appendix 9 or by using proprietary kits. White colonies were selected for further growth and plasmid preparation.

2.7.2 pCR-Script cloning system (Stratagene Inc., USA)

This system allows efficient cloning of unmodified PCR products at relatively high efficiency. The ligation efficiency of blunt-ended DNA fragments is increased by the simultaneous re-cleaving by *Srf* I of nonrecombinant, religated vector DNA. *Srf* I is a novel rare-cleavage restriction enzyme that recognizes the sequence 5'-GCCCGGGC-3', consequently the PCR fragment must be free of these sites and neither primer can have the sequence 5'-GGGC-3' at the 5' end. The pCR-Script cloning vector is based upon the pBluescript II SK(+) phagemid. Epicurian coli XL1-Blue MRF' supercompetent cells are included in the kit. Ligation reactions and transformation of bacteria were carried according to manufacturer's instructions.

2.7.3 TA cloning kit (Invitrogen BV, The Netherlands)

This kit allows the cloning of unmodified PCR products directly into plasmid vector. The method relies on the nontemplate-dependent activity of thermostable polymerases that add a single deoxyadenosine to the 3' ends of the PCR product. These 3' A overhangs complement the single 3' thymidine overhangs in the vector insertion site. Polymerases with 3' to 5' exonuclease activity will give very low cloning efficiencies. Two generations of plasmid were used: pCRII and pCR2.1, the latter differing from the former by the removal of the SP6 promoter site. Ligations were done according to manufacturer's guidelines; PCR products were used as soon as possible after amplification, due to the possibility of degradation of the 3' A overhangs. Transformation was done using the supplied INValphaF' One Shot competent *E.coli*, according to the protocol given.

2.8 Plasmid preparations

A simple and rapid method (2.8.1) was employed to produce DNA suitable for digestion with most restriction enzymes to allow screening of bacterial colonies. The Qiagen miniprep kit (Diagen GmbH) was used (method 2.8.2) to provide plasmid of sufficient purity for use in DNA sequencing, transcription/ translation and use in cell culture. Where larger quantities of plasmid were required the Qiagen midiprep kit was used.

2.8.1 Rapid plasmid preparation

White bacterial colonies were selected using a sterile pipette tip and inoculated into 5 mls of broth (typically LB, appendix 10), with appropriate antibiotic selection, and incubated at 37°C, shaking, overnight. One ml of this culture was pelleted in a 1.5ml microcentrifuge tube and the supernatant discarded. 100μl of TNE (appendix 10) and 100μl of TNE saturated phenol/chloroform/isoamyl alcohol were added then vortexed vigorously for at least 1 min. The phases were separated by centrifuging for 8 mins in a microcentrifuge.

The upper 100μl aqueous phase was removed to a clean tube. The nucleic acid was precipitated with 250μl of ice cold ethanol (the addition of salt in the form of 1/4 volume 10M ammonium acetate was found to be unnecessary). The nucleic acid was pelleted in a microcentrifuge by centrifuging for 15 mins. The supernatant was removed and the pellet air dried. The pellet was washed with 75% ethanol if extra salt had been added in the precipitation. The nucleic acid was dissolved in 70μl of water; 3μl of this was usually sufficient for restriction enzyme digestion (including RNase A).

2.8.2 Qiagen (Diagen, Germany) mini and midi plasmid preparation kits

The procedure is based upon a modified alkaline lysis method and the adsorption of DNA on to silica, held in a column, in the presence of high salt concentration.

The mini-prep system can isolate up to 20μg of plasmid from 1-5mls of bacterial culture grown to saturation. The plasmid is eluted into up to 100μl water or TE (pH 8.0). The procedure was carried out according to manufacturer's guidelines.

To give optimal DNA concentration typically 3 mls of a 5 ml overnight culture was used and the resultant DNA eluted into 70μl water.

The midi-prep system can isolate up to 100μg of high copy number plasmid, typically from 30mls of bacterial culture grown to saturation.

2.9   Construction of HEV cDNA plasmid library

Total RNA from liver (method 2.2) was used as template for generation of cDNA using two primer strategies, either oligo dT or an HEV specific primer containing an EcoRI restriction site. Since HEV was known to have an internal EcoRI site (see restriction endonuclease site map, appendix 8), digestion of the cDNAs with that enzyme would lead to generation of fragments with either EcoRI overhangs at each end, an EcoRI overhang and a blunt end, or blunt ends only (where dT primed synthesis stalled prior to natural EcoRI site). Cloning vectors had to be restricted appropriately.

2.9.1 cDNA synthesis

Two aliquots of total RNA were microfuged at 10,000g for 25 minutes and the resultant pellets air dried and each dissolved in 13.5 µl tissue culture grade water. All the RNA from one of the pellets (13.5 µl) plus 10µg primer (either oligo dT or HEV specific 3' terminus, see set 2, table 2.1) was heated to 70°C for 2 minutes, placed on ice and microfuged briefly. The first strand cDNA synthesis reaction comprised the primed RNA plus 1 mM DTT (Gibco BRL, UK), 1x first strand buffer (Gibco BRL, UK), 1 mM each dNTP, 100 units HPRNI and 1000 units murine reverse transcriptase (Gibco BRL, UK) in a final volume of 50µl. The reaction was incubated at 37°C for 75 minutes before briefly centrifuging and proceeding to the second strand synthesis.

The entire first strand reaction was used in the second strand reaction with addition of 70µl 10mM $MgCl_2$ (filter sterilised), 10µl 1M Tris pH 7.5 (ultrapure), 1.5µl 1M ammonium sulphate (filter sterilised), 1µl RNaseH (1000u/ml), 4.5µl *E. coli* DNA PolI (22.5u) and 5µl water (TC grade) to a final volume of 142µl. The reaction was left at 12°C for 1 hour then 22°C for a further hour, following this the enzymes were heat inactivated at 75°C for 10 minutes and the reaction cooled on ice and briefly microfuged.

In order to generate blunt ends to the cDNA further dNTPs were added (original component destroyed during heat denaturation) to a final concentration of 1mM each, along with 2µl T4 DNA polymerase (10 units). This reaction was left at 37°C for 25 minutes.

32µl of the final reaction containing the blunt ended cDNA was stored at -20°C and the remainder was phenol extracted.

2.9.2   Ligation of cDNA into plasmid cloning vector

Approximately 110µl of each of the phenol extracted cDNAs was fractionated using a Sepharose CL-4B column, allowing passage of fragments over 400 bp, according to manufacturer's criteria (Size-sep columns, Pharmacia Biotech). The columns were equilibrated with *Eco*RI reaction buffer (1x concentration).

The entire eluates were digested using 40 units of *Eco*RI at 37°C for 75 minutes. This reaction mix was fractionated on two new columns equilibrated with 1x T4 DNA

ligase buffer (Gibco BRL, UK). 102μl of digested oligo dT primed cDNA was obtained and 98μl of digested HEV specific primed cDNA.

Ligation to vector (pBluescript KS+) was performed as follows.

Reaction A- 240ng *Eco*RI/ *Hinc*II digested vector (8μl) with the 102μl of *Eco*RI digested oligo dT primed cDNA, and 5μl (5 units) T4 DNA ligase (Gibco BRL, UK);

Reaction B- 150ng *Eco*RI digested and phosphatased vector (5μl) with 50μl of HEV specific primed, *Eco*RI digested cDNA, and 3μl T4 DNA ligase;

Reaction C- 120ng *Eco*RI/ *Hinc*II digested vector (4μl) with 48μl of HEV specific primed, *Eco*RI digested cDNA, and 3μl T4 DNA ligase.

Small volumes of 5x T4 DNA ligase buffer were added to counteract the dilutional effect of the vector and enzyme in the reaction. A re-ligation control was set up using 30ng of *Eco*RI digested and phosphatased vector. All ligations were left at 12°C for around 15 hours then placed at –20°C.

2.9.3 Transformation of bacteria

Around 2 ng vector equivalent (1-2μl each ligation) was used to transform ultra-competent *E. coli* (Stratagene Inc., USA) with efficiencies of between 1-2 x 10^9 transformants/μg. Transformation was carried out according to the manufacturer's protocol. The final volume of transformation mix was calculated to avoid colony overcrowding when divided between plates. LB ampicillin (100μg/ml) agar medium (appendix 10) was used, either in standard or 135mm diameter circular plates, using the blue/white (X-Gal and IPTG) indicator system. Typically, 2μl of ligation A yielded around 30,000 colonies.

The transformation solutions were directly spread onto a nylon membrane (Hybond N, Amersham Intl, UK) (A) and incubated at 37°C overnight. A similarly sized nylon membrane (B), previously wetted on an unused LB ampicillin agar plate, was placed on top of the colonies on membrane A. Efficient transfer was achieved by pressing membranes A and B together with a glass plate, cleaned by flaming between uses. Placing three unequally spaced holes through the membrane sandwich using a sterile needle allowed later orientation of the replica membranes. A further replica, C, was made by similar transfer from membrane A. Membrane A was placed back on to the

original agar plate; membranes B and C were put onto clean plates. All three were incubated at 37°C until colonies around 1 mm appeared on membranes B and C. Plates/membranes A and B were placed at 4°C and membrane C was processed for hybridization screening.

2.9.4 Colony fixing and prehybridization

Prior to screening by hybridization, the bacterial colonies on the nylon membranes were lysed and fixed as follows. Three trays were used containing filter paper wetted with different solutions so that the membranes could be layed on each sequentially:

Tray 1- 0.5M NaOH, membranes soaked for 5 mins and dried on filter paper for 5 mins.

Tray 2- 1M tris pH 7.5 for 2 mins and dried on filter paper for 2 mins.

Tray 3- 0.5M tris pH 7.5/ 1.5M NaCl for 15 mins and dried on filter paper for a minimum 2 mins.

The nucleic acid on the membranes was fixed by baking between filter paper sheets at 80°C for 1.5 hours.

Prehybridization was done at 65°C in 6x SSC, 5x Denhardt's, 0.5% SDS (appendix 10) and 100μg/ml calf-thymus DNA (denatured by boiling) for a minimum of 6 hours. Ten circular membranes, with a blank membrane on top, were placed in 25 mls of prehybridization solution in an appropriately sized petri dish and placed in an outer container inside a shaking incubator.

2.9.5 ^{32}P labelled DNA probe synthesis

Cloned HEV PCR products were used to cover two areas of HEV cDNA that may have been present in the library vector ligations. The PCR set 1 product from position 6282 to 6724 (table 2.1) was chosen to cover sequences upstream of the HEV *Eco*RI site at 5817, and the set A product from 4538 to 4836 (table 2.1) was chosen to cover the remainder.

A commercial kit was used to generate ^{32}P dCTP labelled DNA for use in Southern or northern hybridizations. The Rediprime DNA labelling kit (Amersham Intl, UK) was capable of labelling up to 25ng of DNA to an activity of around 2×10^9 dpm/μg. This

kit was based upon the random oligonucleotide priming procedure described by Feinberg and Vogelstein (1983). Probes were prepared according to the manufacturer's guidelines. The high efficiency of this method meant purification of labelled product was unnecessary. Labelled DNA was denatured by boiling for 5 mins in the presence of the calf-thymus DNA prior to addition to pre-warmed hybridization solution.

2.9.6 Hybridisation and blot washing

Hybridization was carried out overnight in a shaking incubator at 65°C using 15 mls pre-hybridization solution saved from earlier preparation.

Following hybridization, the membranes were washed with 2x SSC at room temperature for 10 mins, then with 2x SSC at 65°C for 15 mins and 2x SSC with 0.1% SDS at 65°C for 15 mins.

Autoradiography was done at -70°C typically for 4 hours in the first instance and repeated depending on the strength of signal.

2.9.7 Selection and screening of potential recombinant plasmids

When the autoradiography result showed signal from a clearly identifiable colony, this was selected and grown in culture as described previously. If an individual colony could not be identified, due to overcrowding of the plate, a representative area was sampled and re-plated onto standard sized LB ampicillin agar. Re-screening of these was either by DNA hybridisation or individual colony plasmid preparation, depending on the colony density. Possible recombinant plasmids were prepared (method 2.8.1) and digested using the restriction enzyme *Pvu*II by standard techniques, to excise the polylinker area of pBluescript (appendix 3). Plasmids containing cloned DNA were processed for nucleotide sequencing (method 2.15).

2.10 Full length genomic HEV RT-PCR

Advances in PCR techniques opened the possibility of amplifying the entire HEV genome in a single PCR reaction. This would simplify the sequencing strategy by allowing subcloning and use of sequence specific sequencing primers, in addition to

facilitating the later expression *in vitro* of HEV. Since this long range PCR required the presence of full length HEV first strand cDNA, an improved reverse transcription method was used.

2.10.1   HEV reverse transcription for long PCR

To maximize the number of full length cDNAs, SuperscriptII M-MLV RNAseH minus reverse transcriptase (Life Technologies Ltd, UK) was used in conjunction with oligo dT primers.

A combination of 5μg of total RNA dissolved in 21.5μl DEPC water and 5μl oligo dT_{12-18} (1mg/ml) was heated to 70°C for 2 mins and then cooled on ice. Then the following were added: 10μl first strand synthesis buffer (supplied), 5μl 0.1M DTT, 2.5μl dNTP's (10mM each), 1μl HPRNI (USB, USA. 130 units/μl) and 5μl (1000 units) SuperscriptII RT. The reaction was incubated at 45°C for 1 hour then stored at - 70°C. Some aliquots were treated with RNAseH (2 units, USB, USA) by adding the enzyme to the completed RT reaction mix and incubating at 37°C for one hour.

2.10.2   Full length PCR

The conditions giving a product of the correct size were as below. PCR was successful for RNaseH treated and untreated cDNA.

1-5μl of the cDNA synthesized in the above reaction was used as the template. The GeneAmp XL PCR kit (Perkin-Elmer, USA) was employed in a reaction containing 200μM each dNTP, 1.25mM magnesium acetate, 0.4μM primers (see table 2.5), buffer supplied by manufacturer, and 3 units of rTth XL polymerase. An in-house buffer giving final reaction concentrations of 25mM tricine/85mM potassium acetate (pH 8.7), 8% glycerol and 2% DMSO was found to be equally effective. Hot start was achieved using ampliwax PCR Gems (Perkin-Elmer, USA) and thermal cycling was carried out in a Perkin-Elmer 9600 machine using the following parameters: an initial denaturation at 94°C 1 min, then 20 rounds of 94°C 15 sec, 68°C 7 min followed by 17 rounds of 94°C 15 sec, 68°C 7 min with autoextension at 15 sec per cycle and a final extension of 10 min at 72°C. 2-10μl of this PCR reaction was amplified in a second, identical reaction to yield a product visible on agarose gel electrophoresis.

PCR was successful for unmodified primers and primers with 5' phosphorylation.

2.11 Cloning of full length genomic HEV RT-PCR products

The presence of proofreading properties (3'-5' exonuclease) in Perkin-Elmer rTth XL enzyme mix (resulting from the small proportion of Vent polymerase included), meant that the TA cloning kit (method 2.7.3) could not be used directly. The primers could not be designed to incorporate useful restriction sites. In order to clone the PCR product two methods were used: addition of 3' A overhangs (kit supplied by Amersham Intl, UK) in conjunction with the TA kit or a blunt end cloning system used (below).

Due to the presence of incorrectly sized products in the the PCR reaction, full length HEV PCR product was purified by agarose gel electrophoresis, excision of the appropriate band, and extraction using silica (QiaexII, method 2.5.2).

Due to the inefficient nature of the ligation and transformation of long lengths of DNA, ultracompetent cells, as well as competent cells supplied with cloning kits, were used when trying to clone entire HEV genomes. The ultracompetent XL2-Blue MRF' *E. coli* (Epicurian coli. Stratagene Inc., USA) used gave a transformation efficiency of around 5×10^9 cfu per microgram of plasmid DNA.

Linear maps showing the orientation of plasmids containing full length HEV sequence are given in appendix 7.

2.11.1   Addition of 3' adenosine to PCR product and TA ligation protocol

The A-tailing kit (Amersham Life Science, UK) was used to modify the extracted PCR product. The kit uses Tth DNA polymerase to preferentially add 3' adenosine residues to the 3' ends of blunt, double-stranded DNA. The PCR product was assumed to have blunt ends due to the presence of a DNA polymerase with 3'-5' exonuclease activity in the reaction. Since agarose gel extracted PCR product was used, the DNA to be modified was added directly to the A-tailing reaction according to manufacturer's guidelines. Approximately 1μg of extracted PCR product was used. This was then purified using Qiaex II (method 2.5.2) and eluted into 23μl of TE pH 8.0.

Two ligation reactions were performed using the pCR II vector supplied in the TA cloning kit. One ligation used a molar ratio of vector to insert of 1:1.5 (pCRII 50 ng, HEV 150 ng), and the other 1:5 (pCRII 25 ng, HEV 250 ng).

2.11.2   Blunt-end cloning system

The General Contractor DNA cloning system (5 Prime 3 Prime Inc, USA) was used to clone unmodified full length HEV PCR products. This system can be used with DNA fragments with a variety of ends- 5' and 3' overhangs, blunt ended, mixed ended, phosphorylated or dephosphorylated. The pCNTR shuttle vector (appendix 6) is only 2.7 kb and is supplied pre-cut and dephosphorylated. The DNA to be cloned can be blunted and phosphorylated prior to ligation using reagents supplied. The kit is supplied with competent *E.coli* and blue/ white colony screening allowed selection of possible recombinant clones.

The full length HEV PCR product used in cloning was derived from HEV cDNA treated with RNAseH. The ligation reaction was done according to the manufacturer's instructions. Approximately 225ng of HEV PCR product was ligated with 150ng of vector.

2.11.3   Transformation of full length HEV PCR product ligations and colony screening

Competent cells supplied with the cloning kits and ultracompetent cells (Stratagene Inc., USA) were transformed with the ligation products from above. Competent cells were transformed with 1μl of the relevant ligation reaction. Colony screening was done using standard methods.

2.12 Determination of the 5' and 3' termini of the HEV genome

Because the extreme 5' and 3' ends of the PCR clones generated are fixed to the primer sequence, an alternative method had to be used to determine their true sequence.

This can be done using Rapid Amplification of cDNA ends (5' and 3' RACE). The 5' RACE method relies upon the homopolymeric tailing of first strand cDNAs and the amplification by PCR of these 5' ends using both a gene specific primer and a primer with a region complementary to the homopolymeric tail. The 3' RACE method is based upon a modified oligo-dT primer first used to synthesize the first strand cDNA and then used as a PCR primer in conjunction with a gene specific primer.

2.12.1 Homopolymeric tailing of cDNA and 5' RACE

One aliquot of first strand cDNA (method 2.10.1) was treated with RNaseH and purified using silica gel (Qiaex II, method 2.5.2) according to the manufacturer's protocol, and eluted into a final volume of 40μl TE pH 8.0 (serial extraction with 20μl twice). 16.5μl was used in each of two homopolymer tailing reactions comprising: 25μM dATP or dGTP, 20 units terminal deoxynucleotide transferase (Promega Ltd, UK) using the buffer supplied by the manufacturer. The reaction was incubated at 37°C for 30 mins. The tailed cDNA was purified using silica gel (QiaexII), eluted into 40μl TE pH 8.0 (serial extraction with 20μl twice).

2.12.2 5' RACE

1μl of each product was used in the PCRs. The following reaction was successful in amplifying the 5' end of the genomic HEV cDNA. GeneAmp XL PCR kit (Perkin Elmer) buffer, 200μM dNTP's, 1.1mM magnesium acetate, 0.2μM primers (tables 2.3 and table 2.1, set 11 positions 516 and 456) and 2 units of rTth polymerase supplied in a 100μl reaction. Hot start cycling using Ampliwax PCR gems (Perkin Elmer) was carried out in a Perkin Elmer 9600 thermal cycler using the following parameters: initial denaturation 94°C 1 min, 37 rounds of 94°C 15 secs, 50°C 20 secs, 72°C 2 min

and a final extension of 10 min at 72°C. The semi-nested round used 2μl of the first round product as target in an identical reaction.

2.12.3   3' RACE

First strand cDNA was made as in method 2.10.1 except that 50 pmoles of 3' RACE oligonucleotide (table 2.3) was used as primer. Following completion of the reaction, the RT enzyme was heat inactivated at 75°C for 5 minutes. 1μl of this (unextracted) cDNA was amplified in a 100μl PCR using 0.3μM primers (table 2.3), 200μM dNTP's, 5% glycerol, 1.5mM magnesium chloride, 500mM KCl, 100mM Tris-HCl pH 8.3, and 2 units Taq DNA polymerase (Advanced Biotechnologies Ltd, UK). Thermal cycling in the PE 9600 was carried out with 94°C 2 min then 38 rounds of 94°C 1 min, 50 °C 30 secs, 72°C 1 min 30 secs followed by a final extension of 10 min at 72°C. The semi-nested round used 2μl of the first round product as target in an identical reaction. Separate aliquots of cDNA were used to generate amplicons for cloning to minimise or highlight any transcription errors introduced by the thermostable polymerase.

2.12.4   Cloning of 5' and 3' RACE PCR products

The PCR products from the RACE reactions were separated by agarose gel electrophoresis and the bands of the correct size were excised. DNA was extracted using QiaexII (method 2.5.2).

The 5' RACE product required A-tailing (method 2.11.1) prior to ligation into the pCRII vector and transformation using the competent cells supplied (method 2.7.3).

Three clones derived from each of the poly-A and poly-G tailed cDNA were selected for plasmid sequencing.

The 3' RACE product did not require A-tailing and a ligation reaction was set up for each PCR product derived from separate cDNA aliquots. Each ligation reaction comprised 20 ng PCR product (4μl), 50 ng pCRII vector (2μl), 4 units T4 DNA ligase (1μl) (USB, USA), 10x ligase buffer supplied (1μl), and 2μl of water, to a total volume of 10μl. The reaction was left at 14°C for 5 hours. Ultracompetent cells

(Stratagene Inc, USA) were used for transformation according to manufacturer's instructions.

Plasmid preparations were made from five clones for use in sequencing; four from one PCR and 1 from a separate PCR.

2.13 Strategy for determining the full length sequence of HEV cDNA

Several methods were used to try to generate a consensus nucleotide sequence for the HEV associated with the case of acute liver failure: multiple short amplicon PCR, cDNA plasmid library construction, full length genomic PCR, and determination of the 5' and 3' termini by RACE. Screening the plasmid library for HEV cDNA was unsuccessful. Three full length HEV clones (A3, B2, GC5- see appendix 7) were used to complete the sequence data. These were either used to generate subclones by restriction digestion (see appendix 8 for positions of useful restriction endonuclease sites) or HEV sequence specific primers (see table 2.4) were used to obtain sequence data from defined points across the full length clone. The majority of the sequence data was obtained for both strands.

Figure 3.5 depicts the contribution of different methods to the consensus HEV sequence.

Target sequence	Primer sequence
Poly A	5' GCTCTAGAAGCTTTTTTTTTTTTTTTTTTT 3'
Poly G	5' GGAATTCGCGGCCGCCCCCCCCCCCCCCCCC 3'
Poly C	5' GGAATTCGGGGGGGGGGGGGGGGGGGGG 3'

Experiment	Primer sequence (5'-3') and 3' HEV position[1]
3' RACE (gene specific)	GCgGattCCaTGGTTAATGT 6653 CTGCTTGAGGATACCATGG 6988
5' RACE (gene specific)	For positions 6104, 6012, 4571and 4279 see table 2.4 For positions 5311, 5282 see set 4 table 2.1 For positions 5608 see set 8 table 2.1 For positions 516 see set 11 table 2.1

Table 2.3: Primers used in 5' and 3' RACE experiments

[1]The numbering is base upon the consensus sequence of the study isolate.

Lower case denotes mismatches with the study strain consensus.

3' HEV position	Strand match	Length (nt)
409	**pos**	20
451	neg	20
742	neg	20
859	**pos**	18
1255	**pos**	18
1300	neg	20
1649	**pos**	20
1895	**pos**	18
1982	neg	20
2028	**pos**	20
2705	**pos**	19
2771	neg	19
2979	**pos**	20
3253	**pos**	19
3711	**pos**	19
3769	neg	20
4279	neg	19
4528 ˙	**pos**	20
4571	neg	19
5420	**pos**	20
6012	neg	18
6081	**pos**	18
6104	neg	21
6434	**pos**	19
6889	neg	19

Table 2.4: Sequencing primers used with full length HEV clones

New primers were designed to completely match the study strain consensus by overlapping with areas already determined. Nucleotide position numbering based on study strain consensus.
Some primers originally used in PCR reactions were also used in sequencing.
Those with positions 4279, 4571, 6012 and 6104 were also used in 5' RACE (see table 2.3).

2.14 Sequence database construction and management

Nucleotide sequences were read by eye from autoradiographs and contiguous sequences constructed using the STADEN version 3 sequence assembly program package (Dear and Staden, 1991).

Database comparisons were done using BLAST (Basic Local Alignment Search Tool, Altschul *et al.,* 1997). BLAST comprises a set of similarity search programs designed to find and rank nucleotide and amino acid sequence matches from a database according to defined statistical parameters.

Multiple sequence alignments were done using version 5.4.1 of the program MULTALIN (Corpet 1988). This was used to highlight differences in the published data for the 5' and 3' NCRs of HEV and to align similar length sequences as required for construction of the phylogenetic trees.

Phylogenetic comparisons were done using the program TREECON version 1.3b (Van de Peer and De Wachter, 1994) using the following parameters: distance estimation according to Jukes and Cantor 1969; insertions and deletions were not taken into account; the tree topology was inferred using neighbour joining; bootstrapping was done at 100 comparisons; and tree rooting used the outgroup option by forced single sequence.

Two phylogenetic comparisons were made. Full length HEV sequences were compared to demonstrate previously described genotype clustering. A short segment of ORF2 was selected for comparison of sequences derived from genotype 1 HEVs. The conserved region chosen was defined by PCR primers described by Wang *et al.,* 1999, selecting 192 nucleotides between 6298 and 6489 based on the Burmese prototype numbering.

2.15 Nucleotide sequencing

Single stranded and plasmid DNA was used in chain terminator sequencing based upon the method of Sanger (Sanger *et al.,* 1977). Sequenase version 2.0 (Amersham Intl, UK) based kits were used throughout, employing ^{35}S dATP and autoradiography.

2.15.1   ssDNA rescue from phagemids

Cloned DNA in pBluescript and pCR-Script (Stratagene inc, USA) (appendices 2 and 3) could be rescued as ssDNA from f bearing bacteria as these vectors contain a filamentous f1 phage intergenic region (M13 related). The host bacteria used were XL1-Blue and XL1-Blue MRF' *E.coli* (Stratagene inc, USA).

ssDNA was generated by first inoculating a single colony from the transformation into 5ml of LB medium with 100µg/ml ampicillin and VCSM13 helper phage (Stratagene inc, USA) at 10^7-10^8 pfu/ml. The culture was grown at 37°C with vigorous aeration for 1-2 hours followed by addition of kanamycin to 70µg/ml (this selects for infected cells). The mix was grown at 37°C for a further 16-24 hours. 1.5mls of cells were spun in a microcentrifuge for 5 mins and 1ml of supernatant removed and added to 150µl of 20% PEG/2.5M NaCl; the bacteriophage DNA was precipitated on ice for 15mins (or 30 mins at 4°C). To obtain a pellet the precipitate was microcentrifuged for 5 mins at 4°C. The supernatant was discarded and the tube re-centrifuged for 30 secs to collect any residual liquid, this was removed and the pellet air dried. The pellet was dissolved in 100µl TE (pH 8.0), extracted with 50µl of phenol and the top 90µl of aqueous phase precipitated using 10µl 3M sodium acetate and 300µl cold ethanol. The DNA was pelleted by microcentrifugation for 15 mins, and the pellet washed with 75% ethanol, re-centrifuged briefly and the supernatant removed. The DNA was air dried and dissolved in 22µl TE (pH 8.0).

2.15.2   Sequencing ssDNA using Sequenase version 2.0 (Amersham Intl, UK)

Standard protocols were used according to manufacturer's instructions and involved annealing template and primer DNA, initial radiolabelling of synthesised DNA and then extension of the DNA synthesis in the presence of dideoxynucleotides (chain terminators).

Typically 0.5-1.0 pmol ssDNA (up to 7µl from pBluescript single stranded DNA preparations, method 2.15.1) were annealed to 0.5pmoles sequencing primer (either M13 forward or reverse, T3, T7 or HEV sequence specific).

^{35}S dATP (1000Ci/mmol) was used as the radiolabel. Initial synthesis of DNA was usually done at room temperature for 3 mins (2-5 mins range). The DNA synthesis

was allowed to continue in the presence of dideoxynucleotides at 37°C for 5 mins. The reaction was stopped with 4µl of stop solution provided. If necessay, reactions were stored at -20°C. The samples were heated to 75-80°C for a minimum of 2 mins prior to loading (2-4µl) on a sequencing gel.

ITP was sometimes substituted for GTP according to manufacturer's instructions. This was done to aid elimination of DNA secondary structure causing difficulty in sequence resolution.

2.15.3   Plasmid sequencing using Sequenase version 2.0 (Amersham Intl, UK) Alkali denaturation and sepharose purification

About 5µg of plasmid (approximately 20µl of Qiagen mini-prep) in 20µl water was used. 5µl of relaxation solution (1M NaOH, 1mM EDTA freshly made) was added, mixed and left at room temperature for 5 mins.

A small spin column was prepared in a 0.5ml microcentrifuge tube by putting two 18 gauge needle holes in the bottom of the tube and overlaying the holes with 10µl glass beads (Sigma, UK, 212-300 micron). 600µl Sepharose CL-6B (Pharmacia, UK), diluted 2 parts sepharose to 1 part water, was added and the tube placed in a larger 1.5ml microcentrifuge tube. The sepharose was packed down and dried by centrifuging in a microfuge for 30 secs, discarding the run-through, and then centrifuging for a further 20 secs. The column was placed in a clean 1.5ml tube, the denatured DNA added to the top of the sepharose column and centrifuged for 1 min. The eluate should be the same volume or slightly less than that applied to the column. The primer was annealed to the denatured DNA as follows: 8.5µl of eluate,1µl of 10x TM (100mM Tris.HCl pH 8.0, 100mM magnesium chloride) and 1µl of primer (10µg/ml) were mixed and incubated at 37°C for 15 mins. This was used immediately in a standard Sequenase sequencing reaction.

2.15.4 Plasmid sequencing using Sequenase version 2.0 (Amersham Intl, UK) Sequenase Quick-Denature Plasmid Sequencing Kit (Amersham Intl, UK)

This kit utilises the property of covalently closed circular DNA to form 'collapsed' structures when denatured, which do not easily re-anneal and have regions which behave like single-stranded DNA. Alkali denaturation appears more efficient than simple heat denaturation.

The kit contains two sets of reagents for denaturing the plasmid DNA. One method involves the addition of glycerol and ethylene glycol to a final concentration of 40% to lower the melting temperature of the plasmid so that denaturation occurs at less than 100°C. The second method involves the addition of sodium hydroxide to denature the plasmid followed by an equimolar amount of hydrochloric acid to neutralize the reaction.

To further improve the sequence quality, 7-deaza-dGTP is used instead of dGTP to reduce the formation of secondary structure in the DNA.

Pyrophosphatase is included to eliminate the possibility of occasional weak bands which can occur with prolonged reaction times.

The standard sequencing protocol is used once primer/template annealing has taken place; typically 0.5pmol of template is used with 2pmol of primer.

2.16 Standard polyacrylamide sequencing gel electrophoresis

To make a 30x40cm, 0.4mm thick, 6% polyacrylamide gel containing 7M urea, 1x TBE, the following were combined in a glass beaker-10.5mls of 40% acrylamide:bis-acrylamide (19:1), 29.4g Urea (ultrapure), 7.0mls 10x TBE and water up to 70mls. The solution was stirred until the urea had dissolved.

Polymerization was achieved by adding 420μl of 10% Ammonium persulphate and 70μl TEMED. The gel mix was poured between two glass plates, held apart at the sides by 0.4mm spacers, which were sealed at the sides and bottom using waterproof tape. A slot for the loading wells was formed at the top using a sharks tooth comb. The glass plates were clamped together using bulldog clips on all sides whilst the gel polymerized.

Gels were typically run in 1xTBE in a tank (Type S2, Life Technologies Ltd, UK) at 60W for 2-8 hours depending upon the area of sequence to be read.

At the end of the run the gel was adhered to 3MM chromatography paper (Whatman) and dried under vacuum at 80°C for 1-2 hours prior to exposure to autoradiography film.

2.16.1 Buffer gradient PAGE

In order to read more sequence from one sequencing gel the buffer in the bottom chamber was altered to 1xTBE, 1M sodium acetate (pH 5.2).

This serves to reduce the spacing between the bands at the bottom of the gel allowing a longer electrophoresis time without loss of sequence data.

2.16.2 Formamide PAGE

Areas of strong secondary structure that result in the compression of bands on sequencing gels can usually be eliminated by the inclusion of 20% formamide in the gel mix and running buffer.

These gels run slightly slower than those without formamide and they require fixing in 20% methanol, 5% acetic acid for 30 mins prior to drying.

2.16.3 Glycerol Tolerant PAGE

When the sequencing enzyme is pre-diluted with 50% glycerol buffer (as in the Quick-Denature plasmid sequencing kit 2.15.4) the concentration of glycerol in the labelling reaction increases from 0.8% to 5%. When the plasmid is denatured using glycerol/ethylene glycol and boiling, the labelling reaction contains an even higher proportion of glycerol. In these situations the Glycerol Tolerant Gel Buffer (Amersham Intl, UK) must be used to prevent sequence distortion in the 300-400 base range. This buffer uses taurine in place of boric acid.

A buffer gradient can also be formed with these gels by the addition of 1M sodium acetate to the lower buffer chamber of the gel tank.

2.17 Transcription and translation *in-vitro* of HEV clones

Full length HEV clones and individual ORF 2 and 3 sequences were to be used in *in-vitro* coupled transcription/ translation reactions. Four of the full length clones were in the pCRII vector which contained the SP6 RNA polymerase promoter sequence. To allow direct comparisons between cell-free translation and proposed cell culture experiments, the same full length HEV clone was used for each. Only the clone in the pCNTR vector (clone GC5) was suitable for subcloning into the pCRII vector for use *in-vitro*, and the pCI-neo mammalian expression vector (Promega Inc., USA).

2.17.1 Subcloning of full length HEV into pCRII vector

The full length HEV sequence was excised from clone GC5 (appendix 7, pCNTR vector) using the restriction endonuclease *Pme*I resulting in a blunt-ended fragment. Subcloning into pCRII vector was achieved by first excising a previously cloned short PCR product from the vector using *Eco*RI, heat denaturing the enzyme, and then blunting the vector ends using Klenow polymerase and 50μM dNTPs. The resulting DNA mixture was separated on an agarose gel and the vector band cut out and purified using Qiaex II (method 2.5.2) prior to dephosphorylation and use in ligation. Standard techniques were used for subcloning and transformation.
This protocol allowed the reading frame and polarity to be correctly aligned for transcription/translation.

2.17.2 PCR generation of ORF 2 and 3 DNA for use in protein expression

In order to study proteins generated *in-vitro* from ORF 2 and 3 these sequences had to be created by PCR due to the lack of appropriate restriction endonuclease sites in the full sequence. PCR primers incorporating SP6 polymerase promoter sequences were used, enabling direct expression of the PCR product in coupled *in-vitro* transcription/translation reactions. Unfortunately, insufficient PCR product for direct use meant that plasmid cloning was eventually required prior to expression.
The full length clone GC5 in the pCNTR vector and clone A3 in the pCRII vector, were used as the targets for PCR.

PCR ORF 3

Around 4ng of the full length clones was used as the target in a 100μl reaction comprising 200μM dNTP's, buffer I (AB Ltd, UK. is 50mM KCl, 10mM TrisCl pH8.3, 1.5mM MgCl$_2$ final concentrations), 5% glycerol, 0.4μM primers (table 2.5) and 2 units Taq polymerase (AB Ltd, UK). Cycling parameters were 95°C, 1 min initial denaturation then 37 rounds of 95°C, 30 sec, 60°C, 30 sec, 72°C, 2 min with a final extension of 10 min at 72°C. The PE 9600 (Perkin-Elmer Inc, USA) thermal cycler was used with thin-walled tubes.

PCR ORF 2

Around 4ng of the full length clones was used as target in a PCR employing the GeneAmp XL PCR kit. 100μl reactions comprised 200μM dNTP's, 0.4μM primers (table 2.5), 1.25mM magnesium acetate, 25mM tricine/ 85mM potassium acetate pH8.7, 8% glycerol, 2% DMSO and 2 units of rTth XL polymerase. Hot start was achieved using AmpliWax PCR Gems (Perkin-Elmer Inc, USA). Cycling parameters were 94°C for 1 min initial denaturation then 37 cycles of 94°C, 15 sec, 58°C, 30 sec, 72°C, 3 min 30 sec with a final extension of 72°C for 10 min. The PE 9600 thermal cycler was used with thin-walled reaction tubes.

2.17.3 Cloning of ORF 2 and 3 PCR products

The pCI-neo mammalian expression vector was digested with *Sma*I using standard conditions and extracted using Qiaex II. The vector ends were dephosphorylated using shrimp alkaline phosphatase, which was inactivated by heating to 65°C.

ORF 2 and 3 PCR products derived from the GC5 clone were phosphorylated using T4 polynucleotide kinase according to standard protocol and the DNA extracted using Qiaex II.

Standard ligations reactions were transformed into ultracompetent cells (Stratagene Inc, USA).

The nucleotide sequences of ORF 2 and 3 were determined to check the validity of their use in the transcription/translation experiments.

Primer	Sequence (5'-3')
FULL LENGTH FORWARD	aGGCAGACCACaTATGTGGTCGATGC
FULL LENGTH REVERSE	CAGGGAGCGCGgAACGCAGAAAAGAG
TnT[1] ORF2 FORWARD	atttaggtgacactatagaatACC**ATG**CGCCCTCGGCCTAT
TnT ORF2 REVERSE	CAGGGAGCGCGgAACGCAGAAAAGAG
TnT ORF3 FORWARD	atttaggtgacactatagaataCTGTGTCGGGTGG**AATGA**
TnT ORF3 REVERSE	TGAGCCGGGGCGACCGCGGTTAGC

Table 2.5: Primers used in full length HEV PCR and ORF2 and 3 PCRs

[1]TnT refers to the coupled transcription/ translation sytem. The TnT positive sense/ forward primers have the SP6 polymerase promoter sequence at the 5' end represented in lower case as mismatches to the study isolate consensus.

Other mismatches with study strain consensus arose during primer design using published sequences.

Bold typeface highlights the start codons for ORF 2 and 3.

Full length HEV genomic PCR primers:
Forward: GGCAGACCACaTATGTGGTCGATGC, nucleotide positions for the study strain consensus were 2-26. RACE identified an additional 5' adenosine.
Reverse: CAGGGAGCGCGgAACGCAGAAAAGAG positions 7193-7168; base mismatches are shown in lower case and arose from the use of published sequences for design.

2.17.4   Standard coupled *in vitro* transcription/translation reactions

In order to study the proteins coded for by the HEV clones available, a coupled transcription/translation *in-vitro* system (TnT kit, Promega Inc. USA) was used. The RNA is generated from plasmid DNA using either SP6, T3 or T7 RNA polymerase and protein synthesis is achieved using a rabbit reticulocyte lysate. The reaction mix allows both of these steps to occur in a single tube, generally resulting in an increased yield of protein in comparison to a standard *in-vitro* rabbit reticulocyte system. The reactions can be modified to include canine pancreatic microsomes (Promega Inc. USA) to provide co-translational protein modification and allow the study of processing events. Protease inhibitors or the protein synthesis inhibitor Puromycin could also be added to analyse possible proteolysis. Proteins were labelled with ^{35}S methionine and separated by SDS-PAGE (method 2.19).

As most of the full length HEV clones were in the TA cloning vector (pCRII, appendix 4) the system using SP6 RNA polymerase was used for these. The full length clone in the pCNTR vector (appendix 6) had to be subcloned into pCRII (method 2.17.1). Standard 50µl reactions contained- 1µg plasmid, 25µl rabbit reticulocyte lysate, 2µl buffer (supplied), 1µl SP6 polymerase (supplied), 1µl amino acids (1 mM, minus methionine), 1µl HPRNI (130 units, USB Inc. USA) and 4µl ^{35}S methionine, translational grade 1000Ci/mmol (Amersham Intl, UK). Incubation was at 30°C for 2 hours. Commonly 25µl reactions were used with 0.5µg of plasmid. Plasmids with antisense orientated HEV sequences were used as negative controls. Reactions were stored at -20°C for protein analysis or -70°C if RNA extraction was considered.

Where canine pancreatic microsomes were used the reactions were 25µl in total containing 0.5µg plasmid, 0.5µl buffer, 0.5µl SP6 polymerase, 0.5µl amino acids minus methionine, 0.5µl HPRNI, 12.5µl rabbit reticulocyte lysate, 2µl ^{35}S methionine and 2.5µl microsomes. Incubation was at 30°C for 90 mins.

Where separate ORF2 and 3 clones were translated, these were in the pCI-neo mammalian vector (Promega Corp. USA, appendix 5) necessitating the use of the T7 bacteriophage RNA polymerase in the coupled transcription/ translation kit. This meant reactions had to be incubated at 30°C for 90 mins. All reactions were separated by SDS-PAGE (method 2.19).

2.17.5 Coupled transcription/translation reactions in the presence of protease inhibitors

The extent and mechanism of HEV protein processing during viral replication was largely uncharacterised. In order to investigate possible protein processing during translation *in vitro*, specific protease inhibitors were added to the standard reactions (see table 2.6) and the proteins generated were compared to those formed by a standard reaction.

Reactions including protease inhibitors were done using the full length HEV clone GC5. A 25μl reaction was used, comprising: 12.5μl rabbit reticulocyte lysate, 1.0μl buffer supplied, 0.5μl SP6 polymerase supplied, 0.5μl amino acids (1 mM, minus methionine), 0.5μl HPRNI (65 units, USB Inc. USA), 2.0μl ^{35}S methionine (Amersham Intl, UK), 1.7μl DNA (0.5μg), and varying volumes of protease inhibitor and water, depending upon the concentration required (see table 2.6). Reactions were incubated for 2 hours at 30°C and then placed at -20°C (10μl was stored at -70°C for later RNA extraction). 6μl was loaded for SDS-PAGE (method 2.19).

2.17.6 Evolution of proteins over time during coupled transcription/translation reactions

To check the evolution of proteins produced from full length HEV clones by coupled transcription/ translation, a reaction was sampled at several time intervals. A 60μl reaction was set up comprising: 30μl rabbit reticulocyte lysate, 2.4μl buffer supplied, 1.2μl SP6 polymerase supplied, 1.2μl amino acids (1 mM, minus methionine), 1.2μl HPRNI (156 units, USB Inc. USA), 4.8μl ^{35}S methionine (Amersham Intl, UK), 3.3μl DNA (1.0μg GC5), and 15.9μl water.

5μl aliquots were taken at 15, 30, 45, 60, 75 and 90 minutes and stored at -20°C. Puromycin was added to the remaining 30μl, to a final concentration of 0.1mM, to inhibit translation. This aliquot was left at 37°C overnight. Samples were analysed by SDS-PAGE (method 2.19).

Inhibitor	Specificity	Stock solution	Starting concentration
Antipain-dihydrochloride	Papain, trypsin, cathepsin A and B	20mg/ml in water	50µg/ml (74µM)
Aprotinin	serine protease	10mg/ml in water	2µg/ml (0.3µM)
Bestatin	Amino-peptidase	1mg/ml in 0.15M NaCl	40µg/ml (130µM)
E-64	Cysteine protease	20mg/ml in 1:1 water/ethanol [v/v]	5µg/ml (14µM)
Leupeptin	Serine and cysteine (including plasmin, trypsin, papain and cathepsin B)	1mg/ml in water	0.5µg/ml (1µM)
Pefabloc SC	Serine proteases including trypsin and chymotrypsin	40mg/ml in water	0.5 mg/ml (2mM)
Pepstatin	Aspartate proteases	1mg/ml in methanol	0.7µg/ml (1µM)
Phosphoramidon	Metallo endopeptidases, specifically thermolysin	20mg/ml in water	5µg/ml (0.009mM)

Table 2.6: Protease inhibitors used in expression *in-vitro*

The starting concentrations given are those present in the final reaction.
Experiments were also run with each of the proteases at a concentration 2.5 times the starting concentration.

2.18 Mammalian cell culture

For mammalian cell culture expression studies, the full length HEV sequence was excised from clone GC5 (appendix 7) using the restriction endonuclease *Pme*I, resulting in a blunt-ended fragment that was ligated into *Sma*I digested, dephosphorylated pCI-neo vector (Promega, USA). Clones were generated in the correct, and reverse, orientation for expression. Standard techniques were used for subcloning. The pCI-neo mammalian expression vector contains the CMV immediate-early enhancer/ promoter region to provide strong, constitutive expression of cloned fragments. The SV40 enhancer and early promoter is placed upstream of the neomycin phosphotransferase gene, expression of which confers resistance against the toxic aminoglycoside G-418. The SV40 origin of replication allows transient episomal replication of the pCI-neo vector in cells expressing the SV40 large T antigen, such as COS1 or 7. A chimeric intron is included to enhance expression of cloned DNA, and the SV40 late polyadenylation signal provides increased stability of transcribed RNA.

COS 7 cells derived from African Green Monkey Kidney were used for transient expression. Procedures for cell maintenance, splitting and freezing are given in appendix 9.

Lipofectin or lipofectamine reagents (Life Technologies Ltd, UK) were used in transfections for attempted transient expression of HEV sequences. Both of these reagents are formulations of cationic lipids which interact spontaneously with DNA to form a lipid-DNA complex. The fusion of the complex with tissue culture cell membranes results in efficient uptake of the DNA.

All procedures were done using aseptic technique in a class 2 cabinet (ICN Flow Ltd, UK).

2.18.1   Transient transfection of adherent COS 7 cells

One 175 cm^2 flask of confluent cells was trypsinized and suspended in 10 mls of complete media comprising Dulbecco's modified Eagle medium (DMEM, Gibco-BRL Ltd, UK) supplemented with 10% foetal calf serum (FCS), 2mM L-glutamine (Imperial chemicals, UK), and 1mM sodium pyruvate (Gibco-BRL, Ltd, UK). The cell density was calculated using a haemocytometer. The cells were diluted to a

density of 1×10^5 per ml in medium (above), and 2 mls of this was placed into each of the required number of 35mm cell culture dishes. These were placed at 37°C in a CO_2 incubator and left overnight (should become 50-80% confluent). Two different techniques were used: Lipofectin and Lipofectamine.

All transfection solutions were made in 12x75mm sterile tubes:

For Lipofectin-

Only full length HEV sequence in the pCI-neo vector was used. Clones with sense and antisense (negative control) insert DNA were used. Three 35mm dishes were used for each clone.

Solution A- 1.5µg (range 1-2µg DNA) DNA diluted in a total of 100µl using reduced serum medium (Opti-mem, Gibco-BRL Ltd, UK) per dish.

Solution B- 10µl of lipofectin reagent (range 2-20µl) plus 90µl of reduced serum medium per dish.

A and B were mixed gently and left at room temp for 12 mins. The cells were washed with 2 mls of reduced serum medium once, then the A plus B mixture was added to 0.8 mls of reduced serum medium and the resulting 1 ml poured onto the cells. The cells were then placed at 37°C in a 5% CO_2 incubator for 6 hours (suggested range 5-24 hours). The media was then changed for complete media and left for a further 48 hours (suggested range 48-72 hours) checking regularly for cell viability. All but 0.5 ml of medium was removed from each dish and the cells scraped off into this and transferred to a sterile 1.5 ml microcentrifuge tube. Cells were immediately processed for extraction of RNA, DNA and proteins using the TRIzol reagent (Gibco BRL, UK) respectively.

For Lipofectamine-

Several experiments were performed: Full length HEV sequence (GC5 in pCI-neo) in addition to cloned separate ORF 2 and 3 and empty vector (negative control) were transfected. Three 35mm dishes were prepared for each.

The method was similar to that for lipofectin except:

Solution A- 1µg DNA in 100µl of reduced serum medium

Solution B- 6µl of lipofectamine plus 94µl reduced serum medium

These were mixed and left at room temperature for 30 mins for the DNA/ lipid complexes to form (suggested range 15-45 min). Washed cells were overlayed with the 1ml transfection mix comprising A+B and 0.8ml reduced serum medium and then

left at 37°C in a CO_2 incubator for 5 hours. The medium was then replaced with complete medium and the dishes incubated overnight. In the morning a further change of complete medium was done. At 53 hours after the start of transfection, one dish was harvested for RNA extraction (suggested range 24-72 hours); at 68 hours one dish was harvested for use in protein blotting (method 2.19.3) and another for use in radioimmunoprecipitation of proteins (method 2.19.4).

Transient transfection systems often require repeated experiments to optimise the ratio of the DNA and cationic lipid reagent used. Such experiments were not possible due to time constraints.

2.19 Analysis of proteins expressed from cloned HEV sequences

Radiolabelled proteins from TnT reactions were analysed directly by SDS PAGE and by western blotting. Proteins from transiently transfected cells grown in culture were analysed by western blotting or radioimmunoprecipitation.

2.19.1   Sample preparation

Samples taken directly from the TnT reactions were boiled for 3 minutes in the presence of 1x sample buffer (appendix 10) prior to loading on the gel.

For the lipofectin transfection, the cells from each dish were scraped into 0.5 ml of media, disaggregated by pipetting up and down, prior to transfer to a 1.5 ml microcentrifuge tube. Cells were pelleted by centrifugation at 13,000 rpm for 2 minutes. The RNA, DNA and proteins were sequentially isolated using TRIzol reagent (method 2.20.1), according to the manufacturer's protocol. However, the protein pellet was air dried for 30 minutes instead of vacuum dried for 5 mins, and subsequently dissolved in 60μl 1% SDS by heating to 90°C (the 50°C suggested was ineffective).

For the lipofectamine transfection, cells harvested for western blotting were washed with 1 ml of sterile PBS and then scraped off into 0.5 ml fresh PBS, briefly centrifuged in a 1.5 ml tube, resuspended in 200μl 1x SDS-PAGE sample buffer and stored at –20°C. Sample preparation for radioimmunoprecipitation is given under the appropriate method.

2.19.2   SDS-PAGE

Proteins were separated on either 7%, 10%, or gradient polyacrylamide gels, length 20 cm and 1.5 mm thick, using the Bio-Rad Protean II xi (Bio-Rad Laboratories, Ca. USA) electrophoresis apparatus. For gradient gels, the resolving component was made by passage of 15% and 6% gel mixes through a gradient maker. The 15% gel comprised the following- 9 mls of 40% acrylamide/bis-acrylamide 19:1, 9 mls of 1M TrisCl pH8.9, 4.8 mls of 50% glycerol/water, 0.24 mls of 10% SDS made up to 24 mls. This was polymerized using 80µl 10% APS and 12µl TEMED. The 6% gel comprised 3.6 mls 40% acrylamide/bis-acrylamide, 9 mls 1M TrisCl pH8.9, 0.24 mls 10% SDS made up to 24 mls. Polymerization was achieved using 150µl 10% APS and 12µl TEMED. After the resolving gel was set, it was either left overnight under a layer of water saturated butanol or a stacking gel of around 2cm was added. Stacking gels were made by combining the following- 2.81 mls 40% acrylamide/bis-acrylamide, 2.81 mls 1M TrisCl pH 6.7 and 0.225 mls 10% SDS made up to 20 mls and polymerized using 188µl 10% APS and 20µl TEMED.

The 7% gel comprised: 8.4 mls of 40% acrylamide/bis-acrylamide 19:1, 18 mls of 1M TrisCl pH8.9, 21.02 mls water, 0.48 mls of 10% SDS, total volume 48 mls. This was polymerized using 160µl 10% APS and 24µl TEMED.

The 10% gel was supplied preformed.

Tris-Glycine buffer (appendix 10) was used for electrophoresis at 25mA through the stacking gel and 45mA through the resolving gel, the run was stopped when the sample buffer dye was at the bottom of the gel.

If autoradiography was to be done the gels were fixed overnight in 50% methanol / 7% acetic acid. Following fixing, the gels were soaked in Amplify solution (Amersham Intl, UK) for 20 mins to give an enhanced signal and then vacuum dried at 80°C for around 2 hours. Autoradiography was done at –70°C for varying times. Western blotting was done as detailed below.

2.19.3   Western blotting

Western blots were done on proteins derived from both TnT reactions and cell culture experiments. The specific antibody was derived from pre-transplant serum taken from the patient with acute hepatitis E infection.

Following electrophoresis, the gel was soaked in transfer buffer (appendix 10) for 25 minutes and the proteins transferred onto nitrocellulose (Amersham Intl, UK) using a semi-dry blotter at 400mA and <25V for one hour. The proteins were stained in 0.5% Ponceau S, 1.0% acetic acid for 2 mins to visualize any transfer and to allow lane marking. This stain was washed off by several sequential rinses with distilled water and 3.0% non-fat milk powder (Marvel, Cadbury-Schweppes, UK) in sterile PBS (appendix 10). Unbound sites on the nitrocellulose were blocked by soaking in 3.0% non-fat milk powder in sterile PBS at 4°C overnight.

The blot was washed 5 times for 5 mins each in PBS-Tween (appendix 10) prior to incubation with pre-transplant serum (primary antibody) diluted 1 in 250 in 75mls sterile PBS, 1% BSA. Incubation was for 2.5 hours at room temperature. Antibody was removed by washing the blot 5 times for 5 min each with PBS-Tween. The blot was then incubated at room temperature for 1 hour 40 minutes in sterile PBS, 1% BSA containing a 1 in 500 dilution of biotinylated anti-human IgG (Amersham Intl, UK). Washing was as before, followed by an incubation for 30 minutes at room temperature in a 1 in 400 dilution of biotinylated-HRP-streptavidin complex in sterile PBS, 1% BSA. Washing was as above with an extra step using PBS alone. Antigen/antibody complexes were detected using the colour development reagents diaminobenzidine and cobalt chloride.

Insufficient quantities of patient serum prevented optimisation of the western blotting protocol. Alternative sera could not be obtained because of the rarity of HEV infection in the UK and the lack of time needed to arrange for samples to be sent from HEV endemic areas.

2.19.4   Radioimmunoprecipitation of cellular proteins

One 35 mm dish of cells from the lipofectamine transfection (method 2.18.1) were washed with DMEM (ICN Flow Ltd, UK) free of methionine, glutamine and serum and then incubated in fresh identical medium for 5.5 hours. Proteins were then radiolabelled by replacing the medium with 0.5 ml of medium containing ^{35}S-methionine at 100μCi / ml as the sole source of methionine. Cells were incubated at 37°C in a CO_2 incubator for 1 hour prior to being scraped off into the medium and pelleted in a microcentrifuge.

The supernatant was transferred to a 1.5 ml microcentrifuge tube containing 100μl 36% PEG 6000, mixed and precipitated at 4°C for a minimum of 2 hours. The PEG precipitate was microcentrifuged for 10 mins and the supernatant discarded. Pellets were resuspended in 100μl of high salt RIP/ 1mM PMSF (phenyl-methyl-sulphonyl-fluoride, protease inhibitor) (appendix 10) and stored at –20°C.

The cell pellet was vortexed in 100μl of high salt RIP / 1mM PMSF, left on ice for 5 mins and then microcentrifuged for 5 min at 4°C. The resulting supernatant was transferred to a clean tube and stored at –20°C.

Immunoprecipitation was done as follows: both the PEG precipitates and cellular protein pellets were thawed and fresh PMSF added to a final concentration of 1mM. Cellular proteins and supernatant proteins were immunoprecipitated. Protein G sepharose (IgG binding, Sigma chemicals Ltd, UK) was washed and resuspended in RIP / PMSF to the original concentration to remove the ethanol and then 20μl of this was added to each tube. 2μl of irrelevant antibody in the form of anti-HBsAg positive control from AUSAB RIA kit was added and binding was allowed to proceed on ice for 30 mins with occasional mixing. This mixture was microcentrifuged for 5 mins and the supernatant transferred to a new tube; 20μl of patient serum (taken pre-liver transplant) was added to the pellets. This serum/ cellular protein mix was left at 4°C for 4 hours (longer if necessary) before adding 50μl Protein G sepharose and leaving on ice for 30 mins with occasional mixing. These tubes were microfuged briefly and the resulting supernatants (120μl) added to 60μl 3x SDS-PAGE sample buffer (final concentration of 1x) and stored at –20°C. Pellets were washed 3 times with 1 ml 0.5M LiCl, 0.1M Tris.Cl pH8.5 and then suspended in 40μl of 1x SDS-PAGE sample buffer.

This method results in both precipitated proteins and total remaining proteins for both cell culture supernatants and cells.

All samples required boiling prior to electrophoresis and precipitated cell protein samples needed brief centrifugation after boiling to remove the sepharose.

Further attempts at radioimmunoprecipitation could not be done due to insufficient amounts of patient serum.

2.20   Detection of HEV RNA

RNA extracted from explanted liver, TnT reactions and from cells grown in culture was analysed by northern blotting in order to attempt to demonstrate the presence of HEV RNA, including any putative sub-genomic message. The total RNA extracted from liver (method 2.2) was stored under alcohol and could be pelleted and dissolved in DEPC water prior to direct use in northern blotting.

2.20.1   RNA extraction using TRIzol reagent

RNA was extracted from TnT reactions and cells grown in culture using TRIzol reagent (Gibco BRL, UK). TRIzol is a mono-phasic solution of phenol and guanidinium isothiocyanate used in a modified extraction procedure based upon the method of Chomczynski and Sacchi (1987). RNA, DNA and proteins can be isolated sequentially.

RNA extraction was performed according to manufacturer's instructions. Typically, for cell culture, cells from 35mm culture dishes were transferred to 1.5 ml microcentrifuge tubes and processed. For TnT reactions only, 10µl of the completed reactions (typically total reaction volume 25µl) was used.

2.20.2   Glyoxylation of RNA

Glyoxylation protects RNA from degradation by RNase in downstream applications. Typically up to 5µg of total RNA in 10.8µl DEPC TE pH7.6 was added to 10.8µl 6M glyoxal, 32µl DMSO, 6µl 0.1M DEPC sodium phosphate. The mixture was incubated at 50°C for 1 hour and transferred to ice. 8µl of loading buffer (appendix 10) was added prior to electrophoresis. The amount of RNA and volumes were adjusted as required.

2.20.3   Agarose electrophoresis of RNA

Electrophoresis was carried out in a midi electrophoresis tank. The tank was made RNase free either by cleaning with RNaseAWAY (Molecular Bio-Products Inc, USA) or washing sequentially with 0.2% SDS, sterile water, absolute ethanol, followed by soaking in 3% hydrogen peroxide for 10 min and several rinses with DEPC water. Gels were made with RNase free agarose and run in 10mM sodium phosphate, DEPC water. Typically electrophoresis was done at 100V/80mA for 2.5 hours with constant recirculation of buffer from cathode to anode.

The amount of RNA loaded is given in the relevant results section.

2.20.4   Transfer of RNA to nylon membrane

Gels were not processed- the RNA was transferred directly to nylon membranes (Hybond-N, Amersham Intl, UK) by capillary action using 20x SSC (appendix) for 16 hours or more. RNA was fixed by baking the membrane at 80°C for 2 hours.

2.20.5   Radiolabelled nucleic acid probing of Northern blot

The Rediprime DNA labelling kit (Amersham Intl, UK, method 2.9.5) was used to generate ^{32}P-dCTP labelled DNA for use in northern hybridizations.

Prior to probing, the RNA on the nylon membrane was de-glyoxylated by heating to 65°C for 10 min in 20mM Tris.Cl pH8.0.

Pre-hybridization and hybridization was done at 42°C in 6x SSC, 5x Denhardt's solution (appendix 10), 0.5% SDS, 50% formamide and calf-thymus DNA at 150µg/ml. 0.2 mls of pre-hybridisation solution was used per cm^2 membrane. The calf-thymus DNA was denatured by boiling for 5 min and snap cooling on ice prior to addition to the pre-warmed solution. Pre-hybridization was done for a minimum of 6 hours. Hybridization with radiolabelled DNA was done overnight with a minimum of 14 hours. The HEV specific probes were derived from the set 1 PCR amplicon (see table 2.1) located in ORF 2, and the set A PCR amplicon (see table 2.1) located in ORF 1.

Following probing, unhybridized nucleic acid was typically removed by washing twice with 2x SSC at room temperature for 15 mins and then 2x SSC/0.1%SDS at

65°C for 10-15 mins. Any variation in washing protocol is given in the relevant results section. Washes of increased stringency were usually unnecessary. Where reprobing of membranes was required, the membrane was first stripped of hybridised nucleic acid by boiling in 0.1% SDS for 5 minutes and allowing to cool to room temperature without drying. Probed membranes were exposed to autoradiography film at –70°C for varying times.

CHAPTER THREE

RESULTS

The full length nucleotide sequence of the HEV strain studied was derived from short amplicon PCR products, full length HEV genomic PCR products and 5' and 3' RACE PCR products.

3.1 Standard length amplicon HEV PCRs (method 2.6)

Several attempts were made to amplify sections of HEV genome by RT-PCR. Most primers were designed to amplify regions of around 300 bp to 500 bp, based upon the Burmese prototype strain sequence.

It was anticipated that around 20 sets of primers would be required to amplify the entire HEV genome in segments. Despite the use of a semi-nested PCR method (sets 1, 2, and 4, table 2.1) and optimisation attempts, those reactions employing primers sourced from India did not give clean PCR products and some failed to amplify products of the expected size. It was unclear whether primer design (for potential primer target sequence mismatch see table 2.1) or manufacture was responsible for suboptimal performance, and it was likely that a combination of these two factors was causative. Additional primers which were designed in our own laboratory and obtained form UK manufacturers generally gave improved results (e.g. set 10, table 2.1).

The PCR reaction parameters which generated HEV amplicons which could be cloned are given in table 2.2 and the relative contribution of these products to the consensus sequence can be seen in figure 3.5. The eight successful reactions covered the following regions of the HEV genome: nucleotide positions 25-456; 1437-2176; 4538-4836; 4798-5282; 5145-5571 (amplicon contained an internal deletion); 6068-6331; 6282-6724; and 6704-7174.

Figure 3.1 shows an example of a semi-nested PCR from set 2 primers (synthesized in India and kindly supplied by S K Panda), highlighting the presence of multiple,

different sized PCR products. Bands of the expected size were excised from agarose gels and the DNA extracted for cloning. The presence of PCR product in the second round negative control suggests that there may have been contaminating nucleic acid in the primers. These bands were of consistently different sizes to those observed in the reactions with cDNA target and, therefore, were thought not to invalidate the experiment.

Primer set 8 (see table 2.1) resulted in a single amplicon of shorter than expected length (not shown). Cloning and sequencing of this amplicon showed it to be HEV sequence with a deletion when compared to the HEV Burmese prototype. The expected amplicon was located between position 5145 and 5571 of HEV with the deletion between 5183 and 5444. This area was sequenced later from alternative amplicons (see discussion 4.1.2). It was unclear whether this finding represented a natural variant sub-population or an artefact of amplification.
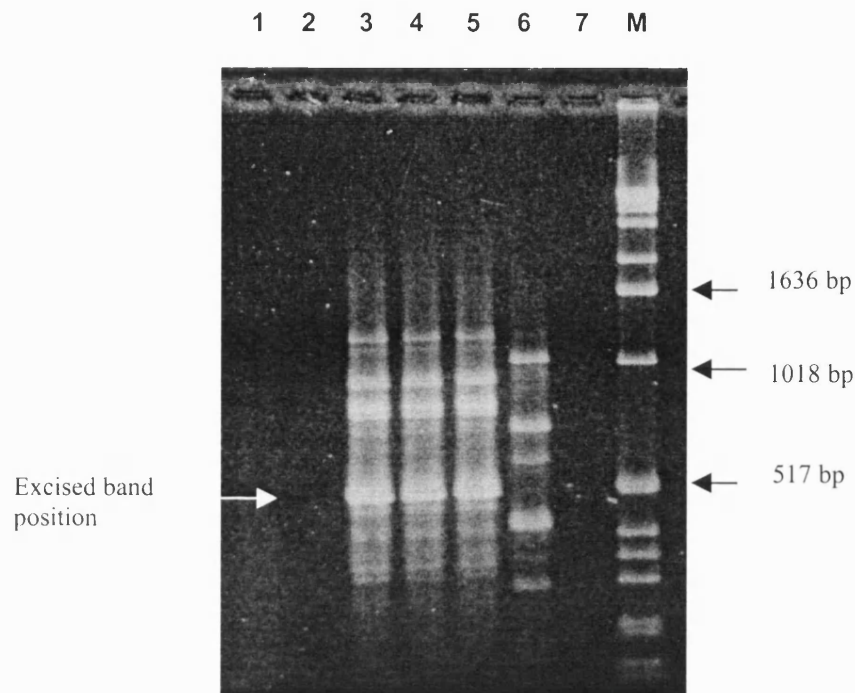
Figure 3.1 Agarose gel electrophoresis of Set 2 HEV PCR products

The set 2 HEV PCR (see tables 2.1 and 2.2) products were separated on a 2% agarose, 1x TBE gel stained with ethidium bromide. The expected correct product size was 500 bp.

The lanes represent (10µl of 100µl reaction loaded for each):
1. First round reaction using 2µl cDNA as target
2. First round negative control (water used as sample)
3, 4, 5. Semi-nested reaction using 5µl of first round reaction as target
6. Semi-nested reaction negative control (first round negative control as sample)
7. Semi-nested reaction negative control (water used as sample)
M. Molecular weight markers (1µg Kb ladder, Gibco BRL, UK), sizes as indicated by arrows on the right-hand side.

3.2 HEV plasmid library

The failure of several primer sets may have represented genuine sequence variations between the strain studied and those published. In order to identify further HEV specific sequence from the total liver RNA, a plasmid based cDNA library was constructed (method 2.9). The integrity of the cDNA was confirmed by using it as target in a successful HEV PCR covering the polymerase gene region (primer set A, see tables 2.1 and 2.2). Following construction of the library, it was validated for presence of plasmid inserts by random colony selection, confirming the presence of cloned cDNA of sizes up to 1kb (not shown), mostly representing ribosomal sequence.

Ligations A (*Eco*RI digested, oligo-dT primed cDNA; *Eco*RI/*Hinc*II digested vector) and C (*Eco*RI digested, HEV specific oligonucleotide primed cDNA; *Eco*RI/*Hinc*II digested vector) gave equal amounts of blue and white colonies; ligation B (*Eco*RI digested, HEV specific oligonucleotide primed cDNA; *Eco*RI digested vector) gave mainly light blue colonies. HEV PCR products spanning positions 4540-4800 (polymerase PCR, upstream of internal *Eco*RI at 5817) and 6264-6725 (set 1 PCR, sequence downstream of internal *Eco*RI) were radiolabelled and used as probes.

Colonies resulting from transformation of all ligations were screened; an estimated 170,000 colonies on plates were probed in total. Many of the identified colonies yielded plasmids without insert. No HEV specific sequeces were identified. Several factors may have accounted for this failure: very low concentrations of HEV specific cDNA as starting material, complicated by the extraction of total RNA rather than polyA RNA; a blunt-ended vector was not used in the ligations which would have facilitated cloning of cDNA fragments synthesised short of the internal *Eco*RI site; and the activity of *Eco*RI used to digest the cDNA may not be optimal at extreme 5' ends.

3.3 Full length HEV genomic PCR and cloning

The introduction of commercially available reagents for long PCR, allowed the use of this technique to amplify full length HEV cDNA. Prior to PCR, an improved reverse transcription method was used to generate suitable first strand cDNA required for the long PCR (see method 2.10.1). The presence of full length cDNA was indicated by successful short amplicon PCR of the extreme 3' and 5' ends of HEV using oligo-dT primed cDNA as target. The full length PCR needed extensive optimisation in order to generate sufficient amplicon. The GeneAmp XL PCR thermostable polymerase (Perkin-Elmer, USA) was used throughout. The original attempts at full length HEV PCR employed the primers covering the extreme 5' and 3' ends of HEV described in table 2.1 (sets 2 and 11), a standard three step temperature reaction, and cDNA synthesised according to method 2.3.2. Despite varying the temperature profile, primer concentration and amount of target cDNA added, amplicon of the correct size could not be generated. Several factors were altered in order to produce a successful reaction, namely the use of improved primers (table 2.5), a different cDNA synthesis method (method 2.10.1), and a two step temperature reaction. The individual contribution of these factors could not be determined. However, correct sized amplicons were not generated each time this method was used, possibly reflecting the inefficiency of the method or the presence of extremely low amounts of full length HEV genome in the original liver tissue or in the extracted RNA. Other important factors included the use of freshly synthesised, RNaseH treated cDNA. The use of extracted and purified amplicon as a target for further PCR was generally inefficient and could not be used to increase product mass for later cloning.

Figure 3.2 shows the successful PCR reaction products separated by agarose gel electrophoresis. The gel shows the second round reaction products including a band of the expected size around 7.2 kb, in addition to larger and smaller bands. The first round reaction did not produce amplicons detectable by gel electrophoresis. The putative full length HEV amplicon was initially confirmed as correct by restriction digestion of the cDNA following extraction from the gel. Other amplicons sized around 200 bp, 2 kb, and between 3-4 kb were also seen in reactions generating the full length amplicon.
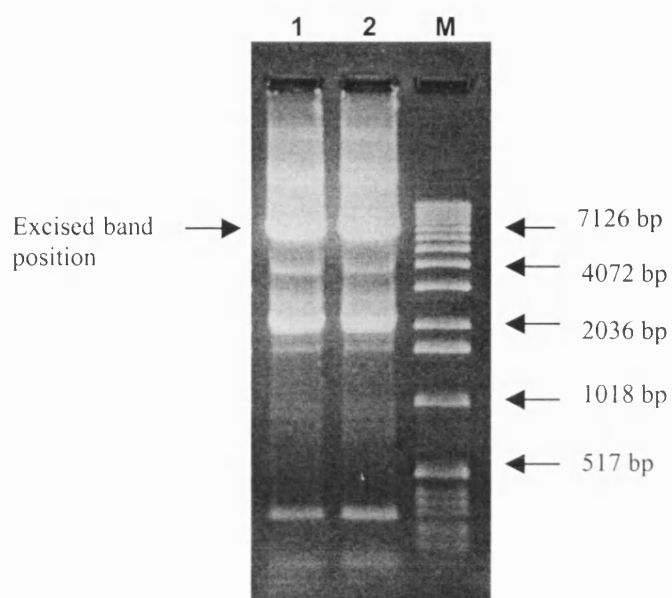
Figure 3.2: Agarose gel electrophoresis of full length HEV genomic PCR products

The products of the full length genomic HEV PCR (see method 2.10.2) were separated on a 1% agarose, 1x TAE gel stained with ethidium bromide. The expected correct product size was 7.2 kb.

The lanes represent:
1. Second round reaction using 2μl of first round reaction as target; 30μl loaded.
2. Second round reaction using10μl of first round reaction as target; 30μl loaded.
M. Molecular weight markers (1μg Kb ladder, Gibco BRL, UK), sizes as indicated by the arrows on the right.

Several different cloning protocols and vectors were used to generate plasmids containing the putative full length HEV sequence. Cloning prior to sequencing was required since there was a relatively small mass of full length PCR product available for purification and direct sequencing, and it was envisaged that cloning into appropriate vectors would allow future protein expression experiments. A low efficiency for the cloning, resulting from poor molar ratios in ligation and high amplicon length, meant that only 6 plasmids containing full length HEV (5 from one PCR reaction coded A3, A9, A20, B2, and B55, and 1 from an independent PCR, coded GC5; see appendix 7) were available for sequencing following screening of around 600 potential transformants. Two plasmids originally grown in bacteria supplied with the TA cloning kit (method 2.7.3) showed low yield, and these were transformed into the bacterial strain XL1 Blue (Stratagene Inc, USA), resulting in a much better yield.

Attempts to purify, clone, and determine the nucleotide sequence of the other amplicons generated by the long PCR, sized between 3-4 kb and at 2 kb, suggested that they were heterogeneous with respect to their content and position in the genome. Therefore, it was unclear whether these fragments represented distinct populations of HEV nucleic acid in the source material or whether they were amplification artefacts.

3.4 Determination of the 5' and 3' termini of the HEV genome

The sequences of the extreme 5' and 3' ends of the HEV cDNA were determined using the RACE method (method 2.12). Several optimisation steps were used prior to generation of PCR amplicons of the expected size.

Figure 3.3 shows the products of the successful 5' RACE experiment separated by agarose gel electrophoresis. Bands of the predicted size were excised from the gel and the DNA purified (method 2.5.2) prior to cloning.

The 3' RACE experiment also required excision of a band of the predicted size from the agarose gel (picture not shown).
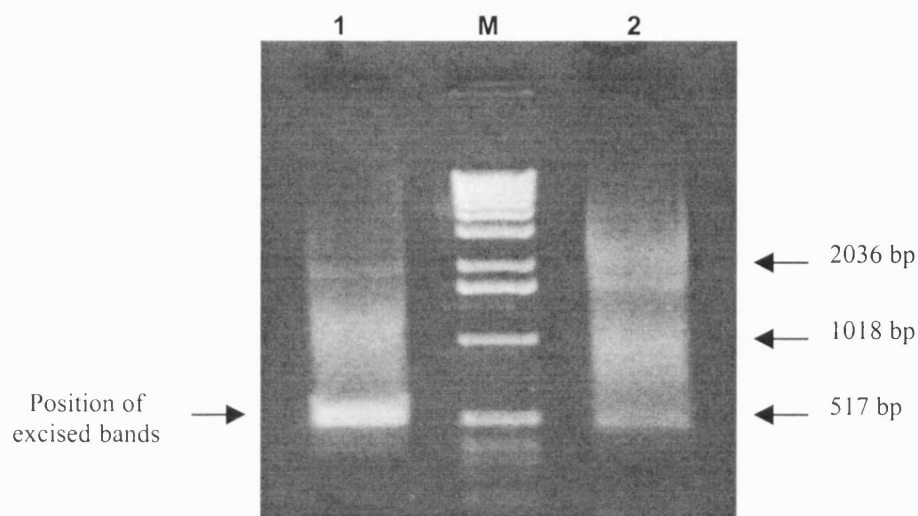
Figure 3.3: Agarose gel electrophoresis of 5' RACE HEV PCR products

The HEV 5' RACE PCR products (see method 2.12) were separated on a 1% agarose, 1x TAE gel stained with ethidium bromide. 2μl of the first round reaction was used as the target in the semi-nested reaction.

The lanes represent (20μl of 100μl PCR reaction was loaded per lane):
1. Semi-nested/ T primer with 'A tailed' cDNA as target
2. Semi-nested/ C primer with 'G tailed' cDNA as target
M. Molecular weight markers (1μg Kb ladder, Gibco BRL), sizes as indicated by arrows on the right side

3.5 Attempts to demonstrate the presence of potential subgenomic RNAs in HEV infected liver tissue

Several attempts were made to detect and define the 5' end of putative previously described (Tam *et al.*, 1991) 3' co-terminal subgenomic RNAs of 2 and 3.7 kb. Northern blot analysis of total liver RNA was performed in order to demonstrate the presence of full length HEV genomic RNA and to define the size of any potential subgenomic HEV RNAs. In addition, the 5' RACE method was used to try to locate the position of any such RNAs.

Northern blot analysis (method 2.20) of total liver RNA (extraction method 2.2) was performed in order to demonstrate the presence of full length HEV genomic RNA and to detect the presence of any subgenomic HEV RNA. Figure 3.4 shows the autoradiographs of the northern blot using the original nucleic acid probe located in ORF 2 and the reprobed blot using probes located in ORF 1 and 2. These probes were chosen to allow identification of possible subgenomic RNA with co-terminal 3' ends. Only bands corresponding to the predicted sizes of ribosomal RNA (4,800 nucleotides and 1,900 nucleotides) were detected.

Primers for the 5' RACE experiments for subgenomic RNA identification were chosen to generate an amplicon of several hundred bases, using a semi-nested PCR reaction (see table 2.3). This assumed the potential subgenomic RNAs were the sizes given above. Multiple PCR reactions gave inconclusive results based on expected product size and the putative subgenomic RNAs were not clearly identified.

In addition to the primers matching the polynucleotide tails, primers at the following HEV 3' positions were used singly or in combination- 7174, 6749, 6724, 6104, 6012, 5608, 5311, 5282, 516, and 456. Three individual cDNA synthesis reactions were carried out for the 5' RACE attempts. The cDNA was tailed with dATP, dGTP and dCTP. The PCR reactions yielded many amplicons, however the majority of these were obtained from only one of the types of tailed cDNAs used as a target. In addition, test restriction enzyme digestion of purified amplicon DNA indicated that they were unlikely to be HEV specific. When the nucleotide sequence of the amplicons was determined, the 5' end was not identical across clones except for one

experiment. This experiment used HEV specific primers located at positions 6104 and 5609 in a semi-nested PCR, resulting in an amplicon of 500 bp from only the dCTP tailed target cDNA. Cloning and sequencing of this 500 bp fragment showed the 5' end to start two nucleotides before the ORF3 start codon at position 5105. This finding was not confirmed using the dATP or dGTP tailed cDNA, and the amplicon was selected from several generated by the PCR.

Original **Reprobed**

1   M   2   3   M   1   2

7460bp ●▶
4400bp ●▶
2370bp ●▶
1350bp ●▶
240bp ●▶
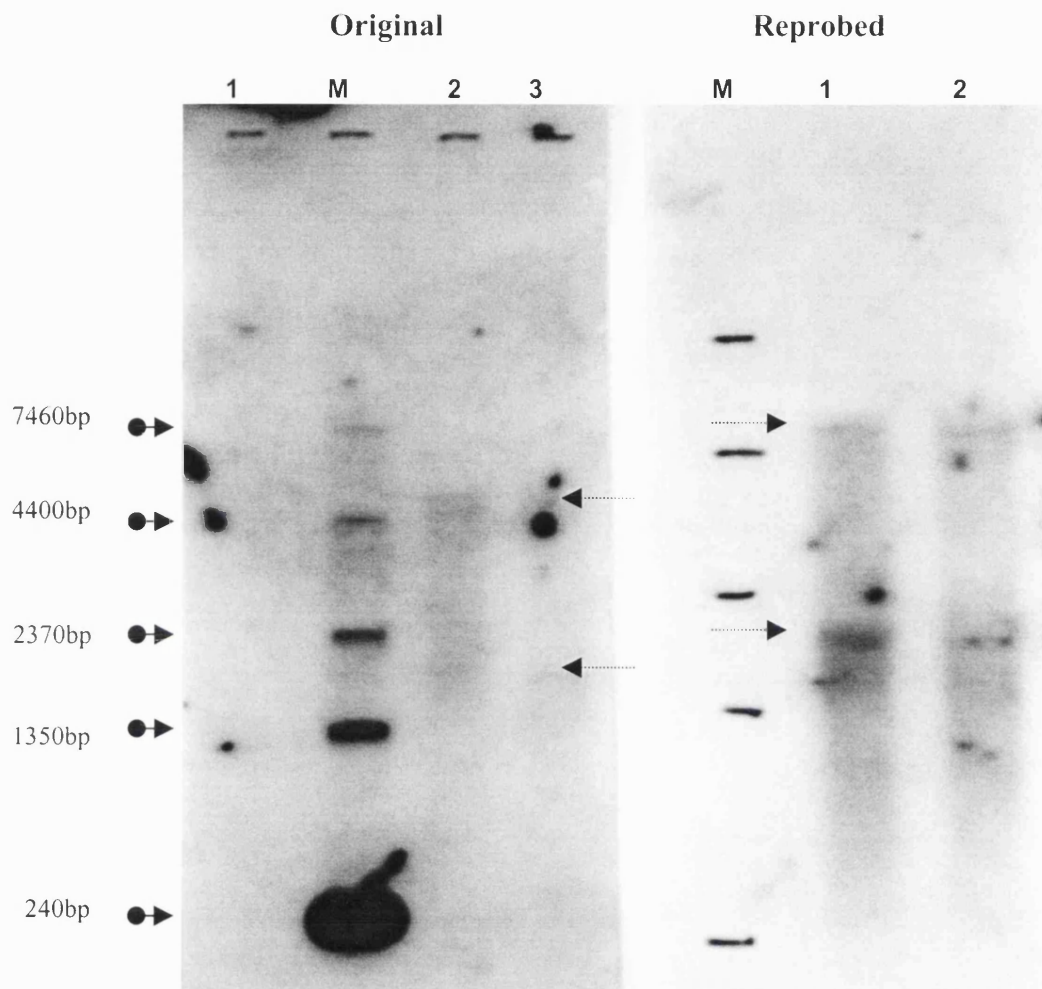
Figure 3.4: Northern blot analysis of total RNA extracted from liver tissue

●▶   Size marker bands; equivalent positions indicated by solid lines in marker lane of reprobed northern blot

┈▶   Indicates bands at approximately 5 kb and 2 kb in both original and reprobed northern blots

The horizontal lines indicating the size marker bands on the reprobed blot were drawn on the blot membrane. The difference in spacing of the size marker positions between the original and reprobed blot is a photographic artefact

Figure 3.4:   Northern blot analysis of total RNA extracted from liver

Original blot

The autoradiograph shows Northern blot analysis of total RNA (method 2.20), using two ^{32}P labelled (method 2.9.5) DNA probes- 25ng lambda bacteriophage (for molecular weight marker detection) and 50ng set 1 PCR product (located in ORF 2, see table 2.1). Autoradiography was for 9 days at -70°C.

The lanes are as follows:
1. 0.6 µg of total liver RNA from primary sclerosing cholangitis (negative control)
M. Poly-A tailed molecular weight markers, (500ng Gibco BRL ssRNA 0.24-9.5 kb)
2. 4.3µg of KS RNA
3. 0.7µg of KS RNA (excess from well 2)

Reprobed Northern blot

The autoradiograph shows the results of re-probing the Northern blot, using three ^{32}P labelled (method 2.9.5) DNA probes- 25ng lamda bacteriophage and 12.5ng set 1 PCR product (located in ORF 2, see table 2.1), 15ng of polymerase gene PCR product (set A, located in ORF1, see table 2.1). Prior to reprobing, the nylon membrane was stripped of hybridised nucleic acid by boiling in 0.1% SDS for 5 mins, followed by cooling to room temperature without allowing drying. Autoradiography was for 11days at -70°C.

The lanes are as follows:
M. Poly-A tailed molecular weight markers (500ng Gibco BRL ssRNA 0.24-9.5 kb).
1. 4.3µg of KS RNA
2. 0.7µg of KS RNA (excess from well 2)

3.6 Sequence construction, analysis and database comparisons

The entire HEV nucleotide sequence was derived from a consensus comprising data from successful short amplicon PCRs, data from 3 full length HEV genome clones (1 from a separate PCR) using HEV-specific sequencing primers and sub-cloning (using *Eco*RI, *Hind*III and *Sac*II as shown in appendix 8), data from 3' and 5' RACE experiments. One full length clone used in expression studies was sequenced fully. The entire nucleotide sequence is given in appendix 1 and has been deposited in GenBank, accession number X98292.

A minimum of 5 clones were used for consensus generation from the short amplicon PCRs and 5' and 3' RACE experiments. Figure 3.5 shows how the sequence data was assembled from the various amplicons generated.

There was little variation between amplicons in comparison to the consensus sequence. For example, the full length clone GC5 contained 24 nucleotide differences when compared to the consensus, comprising 23 transitions and 1 transversion. These changes resulted in 16 amino acid changes, distributed across ORF 1 and 2, with ORF 3 conserved. The relatively high ratio of amino acid changes for the nucleotide changes may indicate that a proportion resulted from artefacts of PCR amplification.
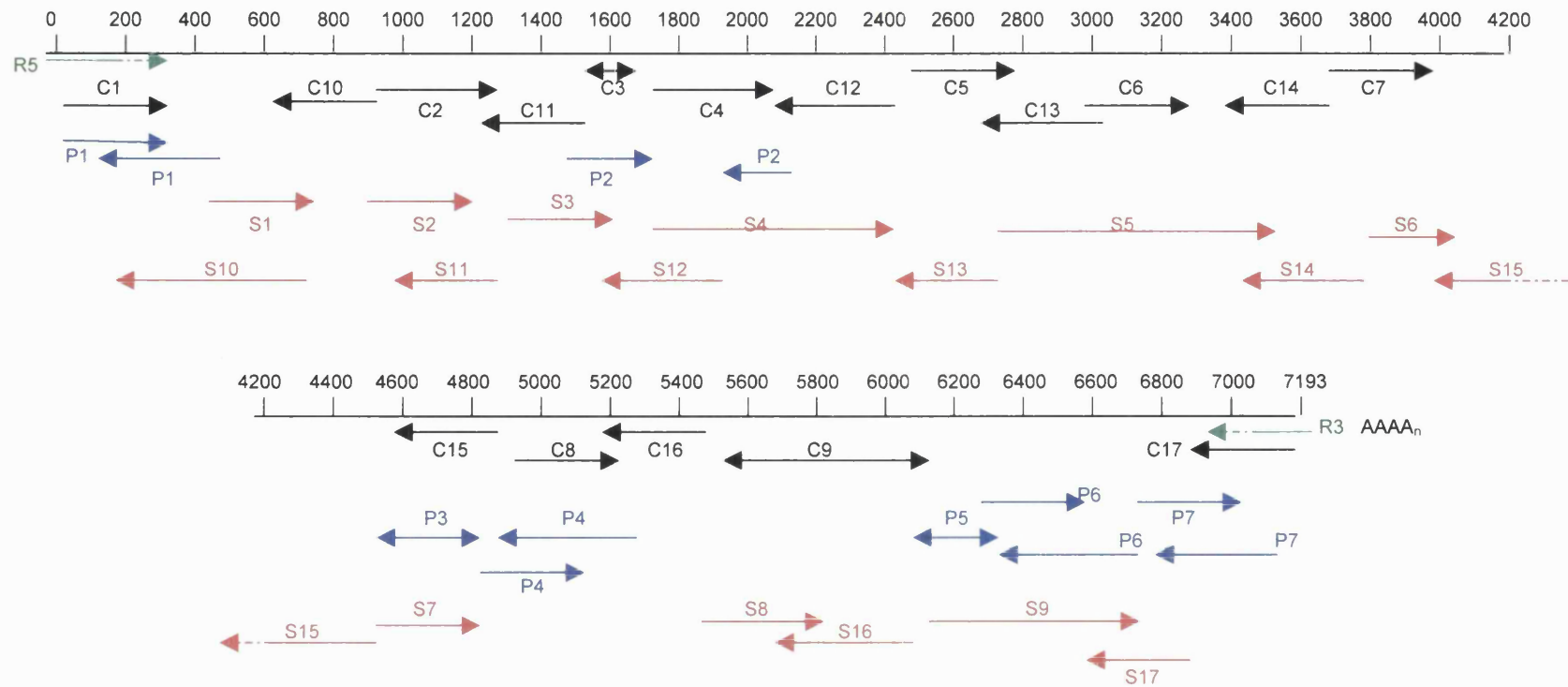
Figure 3.5:   Sequencing strategy for HEV study isolate

Arrows indicate direction and length of sequence; green arrows relate to sequence derived from RACE [R] experiments; black arrows relate to sequence derived from full length HEV clones [C] and sub-clones of full length; blue arrows relate to independent PCR products [P]; red arrows relate to sequence derived from full length HEV clones using HEV specific sequence primers [S]. Primers are not included in the individual sequence lengths.

Figure 3.5: notes

R5- 5' RACE, sequence 0- 250;   R3- 3' RACE, sequence 7193- 7000

C1- Five Full length HEV clones, sequence 26- 340   C17- Five full length clones, sequence 7167- 6872

*Hind*III digestion of full length to yield fragment 1- 4926, *Hinc*II digestion of this fragment to yield C2, C3, C4, C5, C7, C10:

    C2 - Sub-clone 955- 1547, sequence **955- 1252**   C11 - Opposite strand C2 sub-clone, sequence **1547- 1200**
    C3 - Sub-clone **1547- 1697**, both strands sequenced in full
    C4 - Sub-clone 1757- 2419, sequence **1757- 2055**   C12 - Opposite strand C4 sub-clone, sequence **2419- 2066**
    C5 - Sub-clone 2453- 3707, sequence **2453- 2800**   C13 - *Sac*II digest of sub-clone C5 to yield fragment 2453- 3047, sequence **3047- 2700**
    C14 - Opposite strand C5 sub-clone, sequence **3707- 3392**   C6   - *Pst*I digest of sub-clone C5 to yield fragment 2948- 3707, sequence **2948- 3268**
    C7 - Sub-clone 3707- 4926, sequence **3707- 4000**
    C10 - Sub-clone 1- 955, sequence **955- 641**

*Hind*III digestion of full length to yield fragment 1- 4926;   C15 - Sequence **4926- 4600**

*Hind*III digestion of full length to yield fragment 4926- 7193, *Sac*II digestion of this fragment to yield:
    C8 - Sub-clone 4926- 5480, sequence **4926- 5230**   C16 - Opposite strand C8 sub-clone, sequence **5480- 5190**

*Hind*III digestion of full length to yield fragment 4926- 7193, *Sac*II and *Eco*RI digestion of this to yield:
    C9 - Sub-clones **5529- 5817** and **5817- 6144**, both strands of each sequenced in full

P1- Set 11 PCR, sequence **26- 347** and **455- 113**   P2- Set 13 PCR, sequence **1438- 1723** and **2175- 1925** P3- Set A PCR, sequence **4539- 4829** both strands
P4- Set 4 PCR, sequence **4799- 5101** and **5281- 4897**   P5- Set 10 PCR, sequence **6069- 6330** both strands   P6- Set 1 PCR, sequence **6283- 6600** and **6723- 6330**
P7- Set 2 PCR, sequence **6705- 7035** and **7173- 6783**

For HEV specific sequence primers the position is given for the 3' end, p denotes positive sense, n denotes negative sense
S1 - 409p, sequence **430- 753**   S2 - 859p, sequence **884- 1200**   S3 - 1255p, sequence **1274- 1600**
S4 - 1649p+1895p+2028p, sequence **1667- 2418**   S5 - 2705p+2979p+3253p, sequence **2739- 3554**   S6 - 3711p, sequence **3751- 4052**
S7 - 4528p, sequence **4546- 4829**   S8 - 5420p, sequence **5447- 5800**   S9 - 6081p+6434p, sequence **6098- 6785**
S10- 451n+742n, sequence **730- 150**   S11- 1300n, sequence **1285- 956**   S12- 1982n, sequence **1960- 1593**
S13- 2771n, sequence **2761- 2400**   S14- 3769n, sequence **3754- 3392**   S15- 4279n+4571n, sequence **4549- 3961**
S16- 6012n+6104n, sequence **6081- 5750**   S17- 6889n, sequence **6874- 6540**

3.6.1 HEV genome organisation

The genomic structure conformed to the basic layout described for other published HEV sequences, and is described below. The genome length is 7,193 nucleotides, excluding the poly A tail, including a 26 nucleotide 5' NCR and a 3' NCR of 68 nucleotides. ORF 1 begins at nucleotide 27 and comprises 5,079 nucleotides, potentially coding for 1690 amino acids. ORF 2 is located between nucleotide 5,146 and 7,126 with a length of 1980 nucleotides, potentially coding for 660 amino acids. ORF 3 begins at nucleotide 5,105 and ends at 5,473 with a length of 369 nucleotides, potentially coding for 123 amino acids. ORF 3 has a 328 nucleotide overlap with ORF 2 and is translated in the plus 2 frame of ORF1.

Sequence analysis using the BLAST (Altschul *et al.,* 1997) comparison program identified in ORF1 the previously described (Koonin *et al.,* 1992) motifs characteristic of viral methyltransferase, a papain-like protease, RNA helicase (including the NTP binding site represented by amino acids GVPGSGKS) and RNA dependent RNA polymerase (including the amino acids GDD at the active site).

Database comparisons showed the immunodominant regions of ORF 2 and ORF 3 to be well conserved. The study strain had complete identity with the Burmese prototype for the following immunoreactive regions in ORF 2: amino acid positions 25-38, 341-354, and 517-530 (Kaur *et al.,* 1992); 414-433, 442-460 (Khudyakov *et al.,* 1994a, b). For the region 613-654 (identified in Yarbough *et al.,* 1991), residue 614 was different to the Burmese prototype (methionine replacing leucine), however this residue showed variability across genotypes 1 to 4. The study strain also differed from the Burmese protoype at residue 571 within the immunoreactive area 562-580 (Khudyakov *et al.,* 1994b), where the Burmese prototype has a valine but the study strain and representatives of genotypes 2 to 4 have isoleucine. The study strain was conserved against the Burmese prototype for an immunodominant region of ORF 3, amino acids 91-123 (Yarbough *et al.,* 1991).

As expected the hypervariable region, located between nucleotides 2154 and 2384, showed most divergence across strains.

3.6.2 HEV genome comparisons

Genbank was searched using BLAST (Altschul *et al* ., 1997) to identify the closest matches at the nucleotide and amino acid level for the individual ORFs of our HEV strain. The closest nucleotide matches were with genotype 1 viruses, shown below with the reference and accession number:

Full length HEV: L25547 (from Yin *et al,* unpublished) at 93.9%

ORF 1: L25547 at 93.7%

ORF 2: M73218 (from Tam *et al.,* 1991) at 94.1%

ORF 3: M73218 at 97.8%

Many other HEV strains gave similar percentage nucleotide identities.

The closest amino acid match for the putative polyproteins encoded by the separate ORFs were (also see tables 3.3, 3.4, 3.5):

ORF 1: L25547 (from Yin *et al.*), M80581 (from Tsarev *et al.,* 1992) all at 98.4%

ORF 2: M80581 (from Tsarev *et al.,* 1992) at 98.8%

ORF 3: M80581, M73218 (from Tam *et al.,* 1991), D11092 (from Aye *et al.,* 1992) at
    98.4%

Many HEV strains gave similar percentage amino acid identities.

Table 3.1 shows the relative lengths of the genomic regions of HEV for full length representatives of genotypes 1 through 4.

The nucleotide identities for representative full length sequences for each genotype, including our Indian strain are given in table 3.2. The comparisons for the amino acid sequences encoded by individual ORFs derived from published full length sequences are given in tables 3.3, 3.4 and 3.5 (ORF 1, ORF 2, ORF 3 respectively).

Comparisons of the 5' and 3' putative non-coding regions for different genotypes are given in figures 3.6 and 3.7 respectively.

HEV strain	Length (nt)	5'NCR (nt)	ORF 1 (aa)	ORF 2 (aa)	ORF 3 (aa)	3'NCR (nt)
I1	7193	26	1694	660	123	65
I2	7194	27	1694	660	123	65
I3	7194	27	1694	660	123	65
B1	7194	27	1694	660	123	65
B2	7194	27	1694	660	123	65
P1	7163	25	1693	660	123	39
C1	7194	27	1694	660	123	65
C2	7193	26	1694	660	123	65
C3	7194	27	1694	660	123	65
C4	7194	27	1694	660	123	65
M1	7191	24	1692	659	123	74
US1	7186	0	1699	660	123	72
US2	7251	35	1709	660	123	72
T1	7232	25	1708	672/ 658[1]	112	68

Table 3.1: Relative lengths of the genomic regions of full length HEV genotype representatives

The Genbank accession numbers relating to the codes used in the column headings are as follows:
I1- X98292 (Donati 1997); I2- X99441 (von Brunn, unpublished); I3- AF076239; B1-M73218 (Tam 1991); B2- D10330 (Aye 1993); P1- M80581 (Tsarev 1992); C1- D11092 (Aye 1992); C2- L25547 (Yin, unpublished); C3- M94177 (Bi 1993); C4- D11093 (Uchida, unpublished); M1- M74506 (Huang 1992); US1- AF060668 (Schlauder 1998); US2-AF060669 (Schlauder 1998); T1- AJ272108 (Wang 2000).

I1 is the study strain.

The genotypes are assigned as follows:
Genotype 1: B1-2 (Burmese isolates), I1-3 (Indian isolates), P1 (Pakistan isolate);
             C1-4 (Chinese isolates);
Genotype 2: M1 (Mexican isolate);
Genotype 3: US1-2 (United States isolates);
Genotype 4: T1 (Beijing isolate).

[1] Dependent upon start codon usage (Wang *et al.*, 2000)

	I1	I2	I3	B1	B2	P1	C1	C2	C3	C4	M1	US1	US2	T1
I1	*													
I2	92.2	*												
I3	92.0	95.4	*											
B1	93.5	96.1	96.6	*										
B2	93.0	95.4	95.7	98.5	*									
P1	93.9	92.0	92.1	93.6	92.8	*								
C1	93.8	92.3	92.5	93.9	93.4	98.2	*							
C2	94.0	92.0	92.3	93.5	93.0	98.8	98.7	*						
C3	93.8	92.0	92.3	93.7	93.0	98.3	98.2	98.7	*					
C4	93.5	92.7	92.9	94.5	94.1	96.7	97.1	97.2	96.8	*				
M1	75.7	75.0	75.4	75.9	75.7	76.1	75.9	75.9	75.9	75.7	*			
US1	74.4	73.5	73.7	73.9	73.8	74.1	74.2	74.2	74.1	73.7	73.7	*		
US2	74.4	73.8	73.5	74.0	74.0	74.1	74.3	74.3	74.3	73.9	74.5	92.0	*	
T1	74.8	74.8	75.8	75.5	75.5	75.2	77.2	75.3	75.1	75.2	74.5	75.3	76.3	*

Table 3.2: Percentage nucleotide identities of full length HEV genotype representatives

The Genbank accession numbers relating to the codes used in the column and row headings are as follows:

I1- X98292 (Donati 1997); I2- X99441 (von Brunn, unpublished); I3- AF076239; B1- M73218 (Tam 1991); B2- D10330 (Aye 1993); P1- M80581 (Tsarev 1992); C1- D11092 (Aye 1992); C2- L25547 (Yin, unpublished); C3- M94177 (Bi 1993); C4- D11093 (Uchida, unpublished); M1- M74506 (Huang 1992); US1- AF060668 (Schlauder 1998); US2- AF060669 (Schlauder 1998); T1- AJ272108 (Wang 2000).

I1 is the study strain.

The genotypes are assigned as follows:
Genotype 1: B1-2 (Burmese isolates), I1-3 (Indian isolates), P1 (Pakistan isolate); C1-4 (Chinese isolates);
Genotype 2: M1 (Mexican isolate);
Genotype 3: US1-2 (United States isolates);
Genotype 4: T1 (Beijing isolate).

	I1	I2	I3	B1	B2	P1	C1	C2	C3	C4	M1	US1	US2	T1
I1	*													
I2	95.5	*												
I3	96.0	94.7	*											
B1	98.1	96.3	96.9	*										
B2	97.5	95.7	96.2	98.6	*									
P1	98.4	95.9	96.3	98.4	97.7	*								
C1	98.1	95.7	96.1	98.2	97.5	99.2	*							
C2	98.4	95.9	96.3	98.4	97.8	99.6	99.4	*						
C3	96.9	94.4	94.7	96.9	96.1	98.0	97.7	98.1	*					
C4	97.8	95.4	95.8	98.0	97.6	98.9	98.8	99.1	97.4	*				
M1	84.2	81.8	82.2	83.8	83.4	84.0	83.7	83.9	82.8	83.4	*			
US1	82.9	80.7	81.8	82.4	82.3	83.0	82.5	82.8	81.6	82.5	82.0	*		
US2	82.9	80.7	81.3	82.6	82.3	82.9	82.3	82.6	81.6	82.3	82.0	97.5	*	
T1	82.1	80.0	80.6	81.8	81.7	81.3	81.8	81.9	80.7	80.8	79.9	84.4	84.4	*

Table 3.3: Percentage amino acid identities across ORF 1 of full length HEV genotype representatives

The Genbank accession numbers relating to the codes used in the column and row headings are as follows:

I1- X98292 (Donati 1997); I2- X99441 (von Brunn, unpublished); I3- AF076239; B1-M73218 (Tam 1991); B2- D10330 (Aye 1993); P1- M80581 (Tsarev 1992); C1- D11092 (Aye 1992); C2- L25547 (Yin, unpublished); C3- M94177 (Bi 1993); C4- D11093 (Uchida, unpublished); M1- M74506 (Huang 1992); US1- AF060668 (Schlauder 1998); US2-AF060669 (Schlauder 1998); T1- AJ272108 (Wang 2000).

I1 is the study strain.

The genotypes are assigned as follows:
Genotype 1: B1-2 (Burmese isolates), I1-3 (Indian isolates), P1 (Pakistan isolate);
            C1-4 (Chinese isolates);
Genotype 2: M1 (Mexican isolate);
Genotype 3: US1-2 (United States isolates);
Genotype 4: T1 (Beijing isolate).

	I1	I2	I3	B1	B2	P1	C1	C2	C3	C4	M1	US1	US2	T1
I1	*													
I2	98.2	*												
I3	97.1	97.7	*											
B1	98.5	99.1	98.0	*										
B2	97.7	98.3	97.3	98.6	*									
P1	98.8	99.1	98.0	99.4	98.6	*								
C1	98.8	99.1	98.0	99.7	98.9	99.7	*							
C2	98.5	98.8	98.2	99.1	98.3	99.4	99.4	*						
C3	98.2	98.5	98.2	99.1	98.3	99.1	99.1	98.8	*					
C4	98.0	98.4	97.3	98.6	98.0	98.9	98.9	98.6	98.3	*				
M1	93.0	92.9	92.1	93.2	92.4	93.3	93.3	93.0	92.9	92.6	*			
US1	91.1	91.5	90.8	91.8	91.5	91.7	92.0	91.7	91.4	90.9	90.1	*		
US2	91.4	91.8	90.6	91.8	91.5	92.0	92.0	92.0	91.7	91.2	90.6	98.0	*	
T1	89.2	89.2	88.2	89.4	89.1	89.5	89.4	89.5	89.3	88.9	87.7	89.8	91.4	*

Table 3.4: Percentage amino acid identities across ORF 2 of full length HEV genotype representatives

The Genbank accession numbers relating to the codes used in the column and row headings are as follows:
I1- X98292 (Donati 1997); I2- X99441 (von Brunn, unpublished); I3- AF076239; B1-M73218 (Tam 1991); B2- D10330 (Aye 1993); P1- M80581 (Tsarev 1992); C1- D11092 (Aye 1992); C2- L25547 (Yin, unpublished); C3- M94177 (Bi 1993); C4- D11093 (Uchida, unpublished); M1- M74506 (Huang 1992); US1- AF060668 (Schlauder 1998); US2-AF060669 (Schlauder 1998); T1- AJ272108 (Wang 2000).

I1 is the study strain.

The genotypes are assigned as follows:
Genotype 1: B1-2 (Burmese isolates), I1-3 (Indian isolates), P1 (Pakistan isolate); C1-4 (Chinese isolates);
Genotype 2: M1 (Mexican isolate);
Genotype 3: US1-2 (United States isolates);
Genotype 4: T1 (Beijing isolate).

	I1	I2	I3	B1	B2	P1	C1	C2	C3	C4	M1	US1	US2	T1
I1	*													
I2	98.4	*												
I3	94.3	95.9	*											
B1	98.4	100	95.9	*										
B2	96.7	98.4	94.3	98.4	*									
P1	98.4	100	95.9	100	98.4	*								
C1	98.4	100	95.9	100	98.4	100	*							
C2	96.7	98.4	95.9	98.4	96.7	98.4	98.4	*						
C3	96.7	98.4	94.3	98.4	96.7	98.4	98.4	96.7	*					
C4	96.7	98.4	94.3	98.4	96.7	98.4	98.4	97.6	96.7	*				
M1	88.6	87.0	84.6	97.0	87.0	87.0	87.0	85.4	85.4	85.4	*			
US1	84.4	84.4	80.3	84.4	83.6	84.4	84.4	82.8	84.4	82.8	78.7	*		
US2	84.4	84.4	80.3	84.4	83.6	84.4	84.4	82.8	82.8	82.8	80.3	96.7	*	
T1	72.7	72.7	70.9	72.7	72.7	72.7	72.7	73.6	73.4	71.8	71.1	79.1	76.4	*

Table 3.5: Percentage amino acid identities across ORF 3 of full length HEV genotype representatives

The Genbank accession numbers relating to the codes used in the column and row headings are as follows:
I1- X98292 (Donati 1997); I2- X99441 (von Brunn, unpublished); I3- AF076239; B1- M73218 (Tam 1991); B2- D10330 (Aye 1993); P1- M80581 (Tsarev 1992); C1- D11092 (Aye 1992); C2- L25547 (Yin, unpublished); C3- M94177 (Bi 1993); C4- D11093 (Uchida, unpublished); M1- M74506 (Huang 1992); US1- AF060668 (Schlauder 1998); US2- AF060669 (Schlauder 1998); T1- AJ272108 (Wang 2000).

I1 is the study strain.

The genotypes are assigned as follows:
Genotype 1: B1-2 (Burmese isolates), I1-3 (Indian isolates), P1 (Pakistan isolate); C1-4 (Chinese isolates);
Genotype 2: M1 (Mexican isolate);
Genotype 3: US1-2 (United States isolates);
Genotype 4: T1 (Beijing isolate).

```
                        -9              1                                    26
X98292 (I1)        . . . . . . . . . . G GCAGACCAC[C] TATGTGGTCG ATGCC
X99441 (I2)        . . . . . . . AG GCAGACCACA TATGTGGTCG ATGCC
AF076239 (I3)      . . . . . . . AG GCAGACCACA TATGTGGTCG ATGCC
M73218 (B1)        . . . . . . . AG GCAGACCACA TATGTGGTCG ATGCC
D10330 (B2)        . . . . . . . AG GCAGACCACA TATGTGGTCG ATGCC
M80581 (P1)        . . . . . . . . . . GCAGACCACA TATGTGGTCG ATGCC
D11092 (C1)        . . . . . . . AG GCAGACCACA TATGTGGTCG ATGCC
D11093 (C4)        . . . . . . . AG GCAGACCACA TATGTGGTCG ATGCC
M94177 (C3)        . . . . . . . AG GCAGACCACA TATGTGGTCG ATGCC
L25547 (C2)        . . . . . . . . G GCAGACCACA TATGTGGTCG ATGCC
M74506 (M1)        . . . . . . . . . . GCAGACCACA T *TGTGGTCG ATGCC
AF060669 (US2)  TCGACAGGGG GCAGACCACG TATGTGGTCG ATGCC
AJ272108 (T1)      . . . . . . . . . GCAGACCACG TATGTGGTCG A[C]GCC
```

Figure 3.6: Comparison of the 5' NCRs of published full length HEV sequences

The numbering is based upon the study isolate, accession X98292 (I1 above).
No 5' non-coding region has been identified for the HEV strain US1 (AF060668).
[C] Denotes unique nucleotide.
* Denotes space added to highlight alignment.

The Genbank accession numbers relating to the codes used in the row headings are as follows:
I1- X98292 (Donati 1997); I2- X99441 (von Brunn, unpublished); I3- AF076239; B1-
M73218 (Tam 1991); B2- D10330 (Aye 1993); P1- M80581 (Tsarev 1992); C1- D11092
(Aye 1992); C2- L25547 (Yin, unpublished); C3- M94177 (Bi 1993); C4- D11093 (Uchida,
unpublished); M1- M74506 (Huang 1992); US2- AF060669 (Schlauder 1998); T1- AJ272108
(Wang 2000).

The data for sequences P1 and M1 were recently described in Zhang *et al.*, 2001.

The genotypes are assigned as follows:
Genotype 1: B1-2 (Burmese isolates), I1-3 (Indian isolates), P1 (Pakistan isolate);
             C1-4 (Chinese isolates);
Genotype 2: M1 (Mexican isolate);
Genotype 3: US1-2 (United States isolates);
Genotype 4: T1 (Beijing isolate).

```
              1 (7126)                                                     50
X98292  (I1)  TAGTTTATTT GCTTGTGCCC CCCTTCTTTC TGCTGCT... ......TATT
X99441  (I2)  TAGTTTATTT GCTTGTGCCC CCCTTCTTTC TGCTGCT... ......TATT
AF076239(I3)  TAGTTTATTT GCTTGTGCCC CCCTTCTCTC TGTTGCT... ......TATT
M73218  (B1)  TAGTTTATTT GCTTGTGCCC CCCTTCTTTC TGTTGCT... ......TATT
D10330  (B2)  TAGTTTATCT GCTTGTGCCC CCCTTCTTTC TGTTGCT... ......TATT
M80581  (P1)  TAGTTTATTT GCTTGTGCCC CCCTTCTTTC TGTTGCTTAT TT........
D11092  (C1)  TAGTTTATTT GCTTGTGCCC CCCTTCTTTC TGTTGCT... ......TATT
L25547  (C2)  TAGTTTATTT GCTTGTGCCC CCCTTTTTTC TGTTGCT... ......TATT
M94177  (C3)  TAGTTTATTT GCTTGTGCCC CCCTTCTTTC TGTTGCT... ......TATT
D11093  (C4)  TAGTTTATTT GCTTGTGCCC CCCTTCTTTC TGTTGCT... ......TATT
M74506  (M1)  TAGTTTATTT GGCTGTGCCC ACCTACTTAT ATCTGCTGAT TTCCTTTATT
AF060668(US1) TAATTAATTC CTTTTGTGCC CCCTTCGCAG TTCTCTT... TGGCTTTATT
AF060669(US2) TAATTAATTC CTTCTGTGCC CCCTTCGTAG TTTCTTT... CGCTTTTATT
AJ272108(T1)  TGATTTATTT GCTTGTGCCC CCCTTCTGCC ATGTTTT... ...CTTTATT
Consensus     TAgTTtATTt gcTTgtgcCC CCCTTCtttc tgtTgcT... ......TATT

              51 (7167)                                                   100
X98292        TCTCTTTTCT GCGTTTCGCG CTCCCTGAAA AAAAAA.... ..........
X99441        TCTCTTTTCT GCGTTTCGCG CTCCCTGAAA AAAAAA.... ..........
AF076239      TCTCATTTCT GCGTTCCGCG CTCCCTG... .......... ..........
M73218        TCTCATTTCT GCGTTCCGCG CTCCCTGAAA AAAAAAAAA .........
D10330        TCTTATTTCT GCGTTCCGCG CTCCCTG... .......... ..........
M80581        .......... .......... .......... .......... ..........
D11092        TCTCTTTTCT GCGTTCCGCG CTCCCTGAAA AAAAAAAAA .........
L25547        TCTCTTTTCT GCGTTTCGCG CTCCCTG... .......... ..........
M94177        TCTCATTTCT GCGTTCCGCG CTCCCTGAAA AAA....... ..........
D11093        TCTCTTTTCT GCGTTCCGCG CTCCCTG... .......... ..........
M74506        TCCTTTTTCT CGGTCCCGCG CTCCCTGAAA AAAAAA... ..........
AF060668      TCTCATTTCT GCTTTCCGCG CTCCCTGGAA AAAAAAAAA AAAA......
AF060669      TCTTATTTCT GCTTTCCGCG CTCCCTGGAA AAAAAAAAA AAAAAAAAAA
AJ272108      TCCCTTTTCT GCGTCCCGCG CTCCCTG... .......... ..........
Consensus     TCtctTTTCT gcgTtcCGCG CTCCCTG.aa aaa....... ..........
```

Figure 3.7: Comparison of the 3' NCRs of published full length HEV sequences

The numbering in brackets is based upon the study isolate, accession X98292 (I1 above).
For Genbank accession number references and genotype key see legend for table 3.2
For the relative lengths of the 3' NCRs, see table 3.1

3.6.3 Phylogenetic analysis of the HEV sequence

The data generated was used in two separate phylogenetic comparisons. In both cases, HEV sequences for comparison identified by BLAST (Altschul *et al.*, 1997) searches were aligned using MULTALIN (Corpet 1988).

The full length study strain HEV sequence was compared with other full length HEV sequences representing the human isolates of the different genotypes currently described (listed at the foot of table 3.1). Figure 3.8 shows the phylogenetic tree for full length HEV genotype representatives, confirming the genotype clustering as expected.

In addition, a 192 bp section located at the carboxy-terminal end of ORF2 (position 6298 to 6489 of the Burmese prototype) was chosen for comparison of representatives within genotype 1 only. This area was based upon that defined by published ORF2 primers designed to amplify a conserved region across genotypes 1, 2, and 3 (Wang *et al.*, 1999). Incomplete sequences highlighted by the alignment were removed prior to phylogenetic analysis, leaving 27 in total. Figure 3.9 shows the phylogenetic tree for this relatively conserved region of ORF2 for sequences assigned as genotype 1. The genbank accession numbers and the geographical source of HEV (where described) are given below in alphanumerical order in the tree legend.
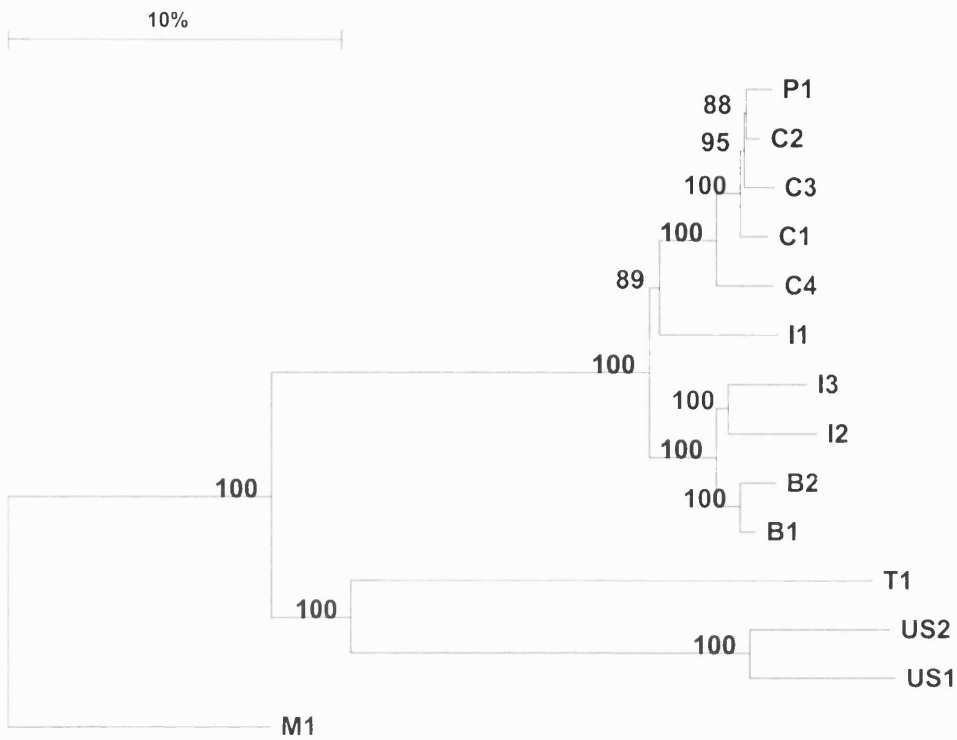
Figure 3.8: Phylogenetic tree of full-length HEV genotype representatives

The tree was constructed as described in method 2.14 with the branch lengths proportional to genetic distance. The bootstrap values (of maximum 100) are given at the nodes. The HEV strain codes and their Genbank accession numbers and geographical origin are identical to those used in table 3.1. I1 is the study strain.

The Mexican M1 sequence (Huang *et al.*, 1992; M74506) was chosen as the outgroup for tree rooting.
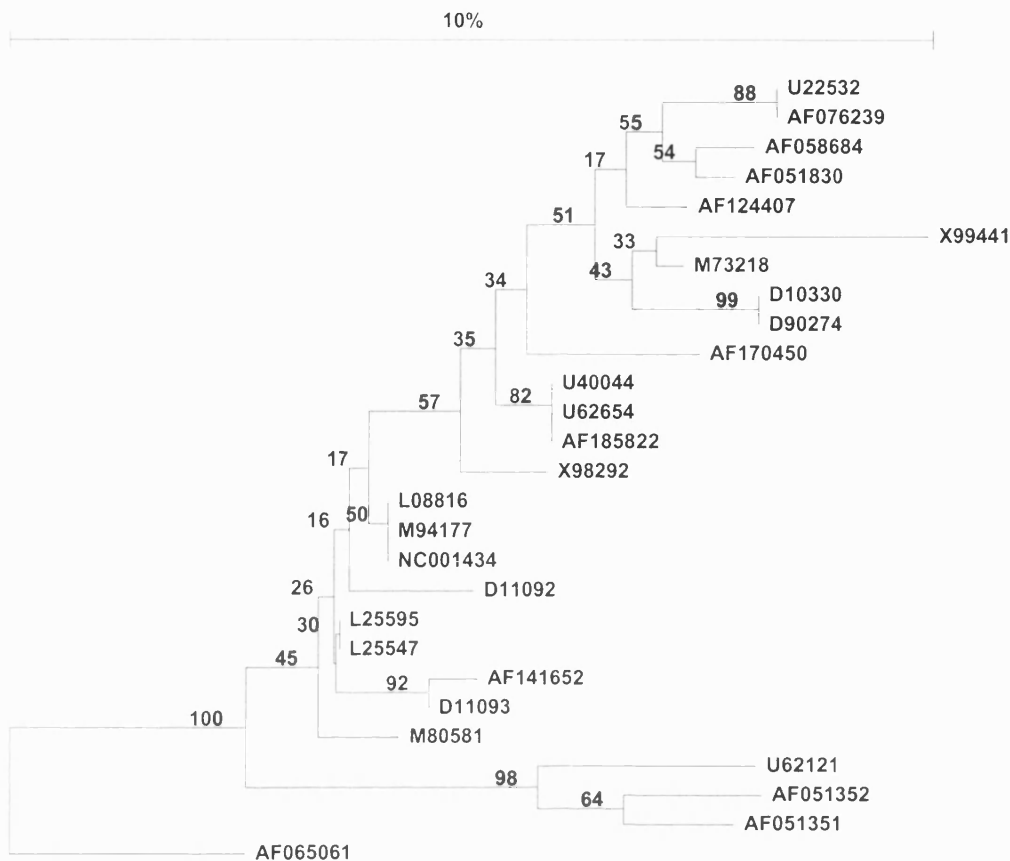
Figure 3.9: Phylogenetic tree of partial ORF 2 sequences from HEV genotype 1 representatives

The tree was constructed as described in method 2.14 with the branch lengths proportional to genetic distance. The bootstrap values (of maximum 100) are given at the nodes. The Moroccan sequence (AF065061) was chosen as the outgroup for tree rooting. The Genbank accession numbers and geographical origin/ strain, where available, are given below in alphanumerical order, including in brackets the coding used for full-length sequences if appropriate (see table 3.1).

AF051351 Egypt, Cairo; AF051352 Egypt, Cairo; AF051830 Nepal, TK15/92 strain; AF058684 Spain, BCN strain; AF065061 Morocco, HEV-Morocco strain; AF076239 India, Hyderabad (I3); AF124407 India, AKL-90; AF141652 China; AF170450 Vietnam; AF185822 Pakistan, Abb-2B strain; D10330 Burma (B2); D11092 China (C1); D11093 China, Uigh179 strain; D90274 origin not described; L08816 China, Xinjiang epidemic; L25547 China; L25595 China (C2); M73218 Burma, prototype strain (B1); M80581 Pakistan, strain SAR-55 (P1); M94177 China, HeBei strain (C3); NC001434 origin not described; U62121 Chad, Chad T3 strain; U62654 Chad, Chad T5 strain; U40044 Pakistan, 2B Abbottabad strain; U22532 India, Hyderabad; X98292 study strain (I1); X99441 India, Madras (I2).

3.7 Expression studies

Subcloning of full length HEV sequence (method 2.17.1) and the separate open reading frames 2 and 3 (generated by PCR, method 2.17.2; cloning method 2.17.3) into appropriate vectors allowed translation experiments both *in-vitro* and *in-vivo*. Western blotting and radioimmunoprecipitation were used to try to demonstrate the presence of HEV specific proteins in mammalian cell culture and from expression *in vitro*. To further investigate the replicative strategy of HEV, northern blot analysis was performed on the RNA extracted from tissue culture and translation experiments.

3.7.1 Expression *in-vitro*

Several experiments were performed to investigate the nature of the proteins produced by cloned HEV nucleic acid. The full length HEV clone GC5 (appendix 7) was used for expression and as a template for ORF 2 and 3 generation, since this clone had been fully sequenced.

Figure 3.10, shows the products of *in-vitro* expression (method 2.17.4) of full length HEV clones A3 and GC5, separated by polyacrylamide gel electrophoresis. The stacking gel artefact can be seen as apparent bands over 200 kDa in lanes 1, 3, and 4. The control experiment in lane 2 shows the presence of a non-HEV derived protein sized between 43 and 68 kDa, common to all experiments.

Translation of full length clone GC5 in lane 3 shows two distinct proteins between 97 and 200 kDa, a protein of around 68 kDa and two further bands between 43 and 29 kDa. Smaller protein bands were less clearly defined. Translation of this clone in the presence of canine pancreatic microsomes appeared to have a deterimental effect on reaction efficiency, with no shift in size of the visible proteins.

Lane 1 shows the translation products from expression of a separate full length HEV clone A3. In comparison to the products from clone GC5, there is absence of the two high molecular weight proteins and the addition of a protein between 68 and 97 kDa. This probably reflects the presence of a premature stop codon in ORF 1 of clone A3 (see discussion 4.2.1).
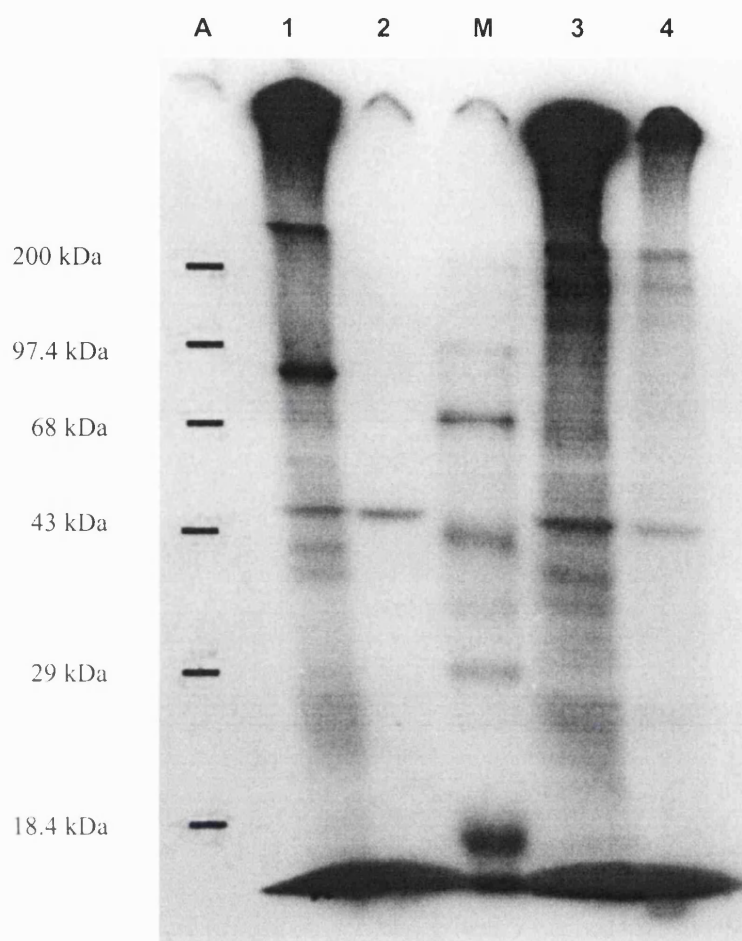
Figure 3.10: Expression *in-vitro* of full length HEV clones

Two full length HEV clones (in the pCRII vector, appendix 4) were expressed *in-vitro* using the Promega TnT system (method 2.17.4, SP6 polymerase version) and the proteins separated by SDS-PAGE using a 10% poyacrylamide, non-gradient stacking gel. Autoradiography was for 2 days at room temperature.

The lanes are as follows:
1. Clone A3 (1µg in a 50µl reaction) 6µl loaded
2. Clone GC5 control, reverse orientation, (0.5µg in a 25µl reaction) 6µl loaded
3. Clone GC5 (0.5µg in a 25µl reaction) 6µl loaded
4. Clone GC5 (0.5µg in a 25µl reaction) translated in the presence of canine pancreatic microsomes, 6µl loaded
M. ^{14}C molecular weight markers (Gibco BRL, UK) sizes as indicated
A. Empty lane. The horizontal black lines are drawn on the autoradiograph to highlight the relative positions of the size marker bands seen in lane **M**

Figure 3.11 shows the comparison of the products of translating a full length HEV clone (GC5 in pCI-neo vector) with those from translation of cloned individual ORF 2 and 3 (method 2.17.4). Multiple protein bands are visible in both the full length expression and expression of ORF 2, however some common proteins are evident. Notably, a proportionately large amount of a protein of around 68 kDa can be seen from ORF 2 expression, and this band is also present in the full length HEV expression products (arrow A). ORF 3 expression appears to have produced a unique HEV specific protein of around 14 kDa, that is not seen in the vector control. Another ORF 3 protein between 14 and 18 kDa may also be present in the full length HEV expression products. The addition of canine pancreatic microsomes has reduced the reaction efficiency without detectable affect on the protein size. Several proteins produced by the vector control expression can be seen in the experimental lanes.
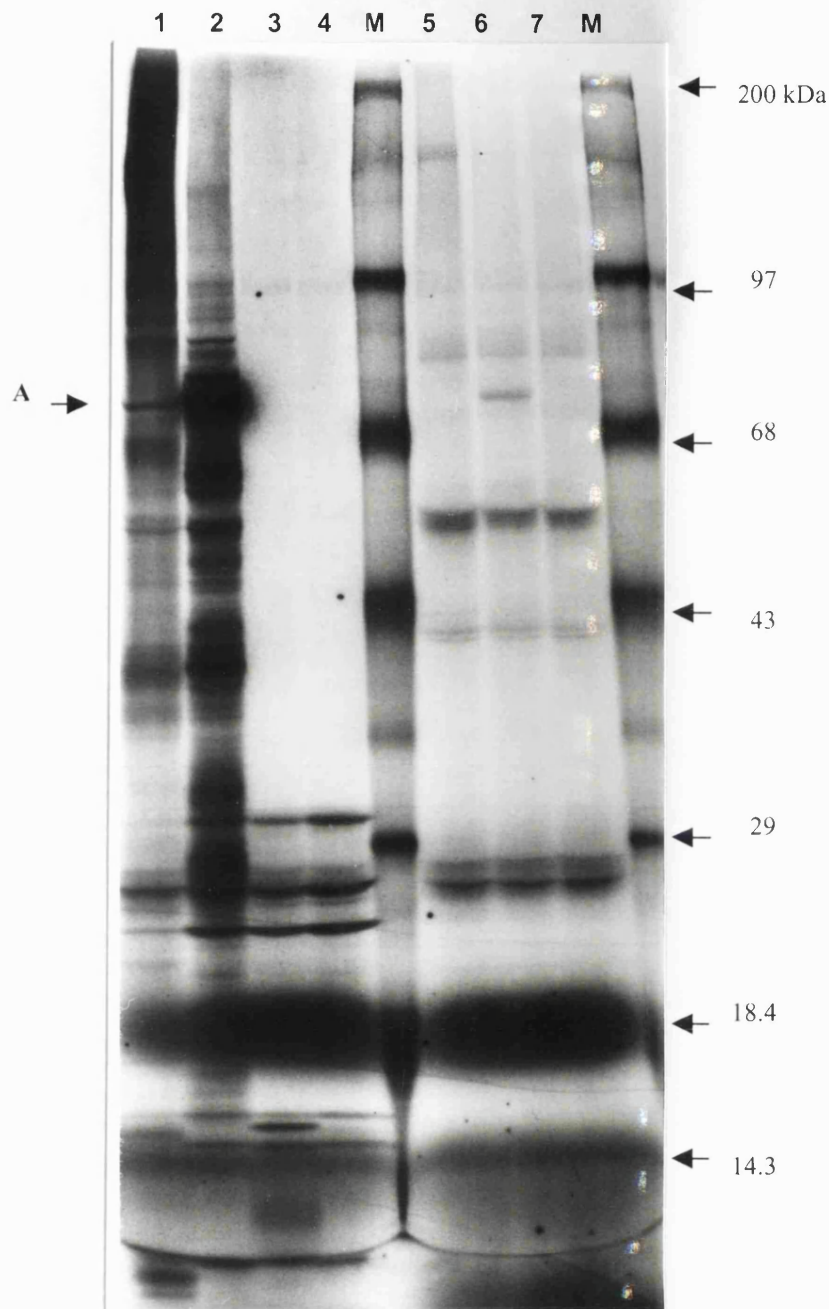
Figure 3.11: Expression *in-vitro* of full length, ORF 2 and ORF 3 HEV clones

A full length HEV clone and cloned individual ORF 2 and ORF 3 were expressed *in-vitro* using Promega TnT T7 version (method 2.17.4). The reactions were 25µl with 0.5µg of plasmid. 8µl of the standard reaction and 10µl of the microsome reaction were loaded on the SDS PAGE gradient stacking gel (method 2.19.2) for protein separation and autoradiography. See text for description of the band shown by the arrow on the left-hand side.

The lanes represent:

M ^{14}C molecular weight markers (Gibco BRL, UK), sizes indicated on the right-hand side

1. Expression of full length HEV genome; **2.** Expression of ORF 2;
3. Expression of ORF 3; **4.** Negative (vector control);
5. Expression of full length HEV genome with microsomes;
6. Expression of ORF 2 with microsomes **7.** Expression of ORF 3 with microsomes

To investigate whether the production of multiple proteins from *in-vitro* expression of cloned full length HEV genome results from proteolytic cleavage, natural degradation or internal initiation of translation, two additional experiments were performed; translation in the presence of specific protease inhibitors, and translation with reaction sampling at short time intervals.

Figure 3.12 shows the products of translation *in-vitro* of cloned full length HEV genome in the presence of different protease inhibitors. The inhibitors were chosen to block a wide range of proteases including serine, cysteine, aspartate, and metallo-endopeptidases (table 2.6). No clear differences can be seen between the pattern of proteins from the positive control and the translations in the presence of protease inhibitors, therefore protease activity was not detected under such conditions. However, the reaction efficiency has been reduced by the protease inhibitors, making comparisons between the different types difficult.

Figure 3.13 shows the shows the time scale of expression of proteins from cloned full length HEV genomes. The positive control protein pattern was as expected from previous experiments. Relatively little translation is evident at time 15 minutes and no clear evolution of proteins can be seen with time, with the maximum amount present at time 45 minutes. Overnight incubation appears to have led to degradation of the proteins, shown by the concomitant decrease in the amount of larger proteins and increase in the amount of smaller products (approximately 14 kDa size range).
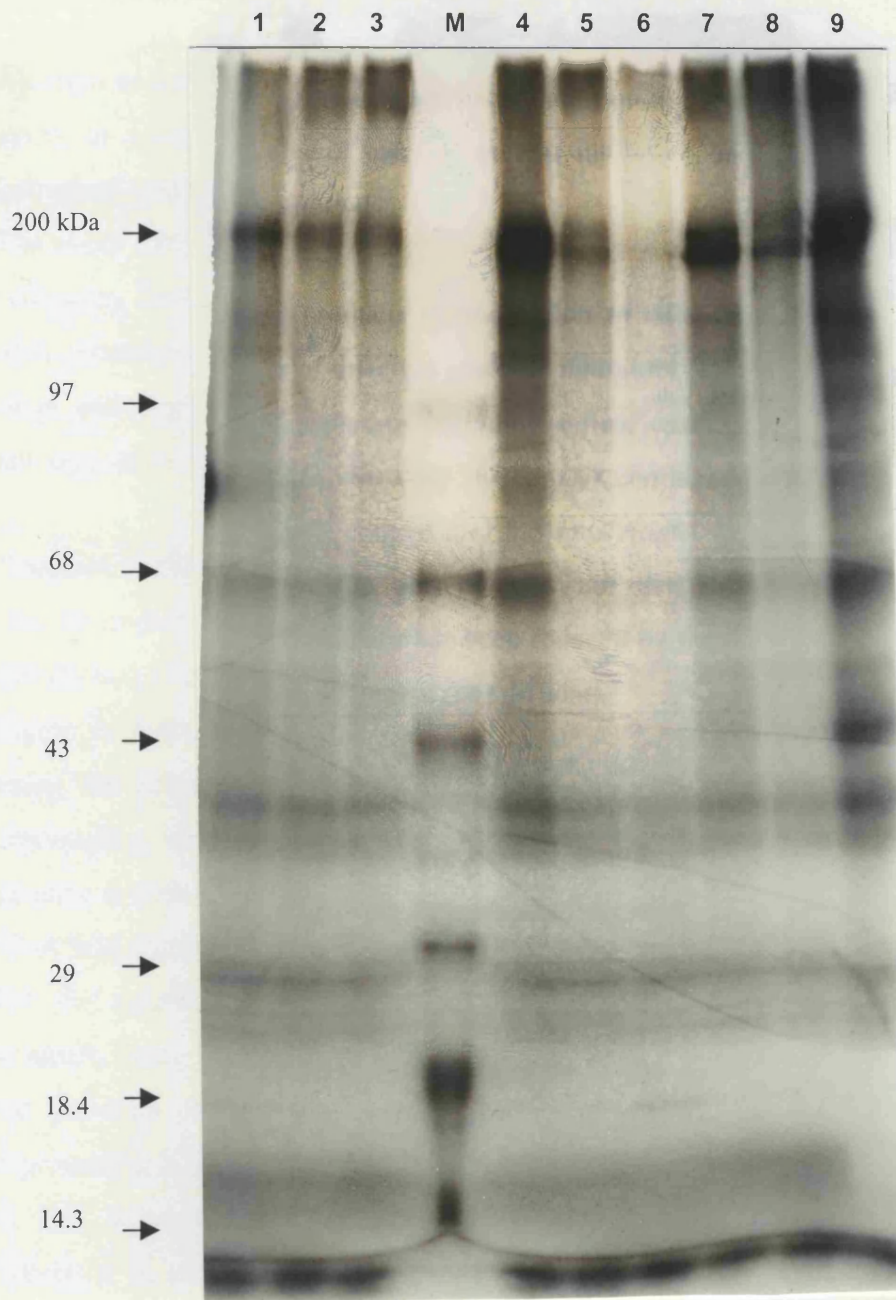
Figure 3.12: Expression *in-vitro* of full length HEV clone in the presence of protease
inhibitors

The full length HEV clone GC5 in the TA vector was expressed *in-vitro* using the Promega
TnT system (SP6 polymerase version, method 2.17.5) in the presence of selected protease
inhibitors (for final concentration see table 2.6). All reactions used 0.5µg plasmid in a 25µl
reaction, with 6µl loaded on the gel. The proteins were separted by SDS-PAGE using a
gradient stacking gel. Autoradiography was for 1 day at -70°C.

The lanes represent the proteases used:

1. Antipain 2. Aprotinin 3. E64   4. Bestatin 5. Leupeptin
6. Pefabloc 7. Pepstatin 8. Phosphoramidon   9. Positive control (no protease)
M. ^{14}C molecular weight markers (Gibco BRL, UK), sizes as indicated by arrows
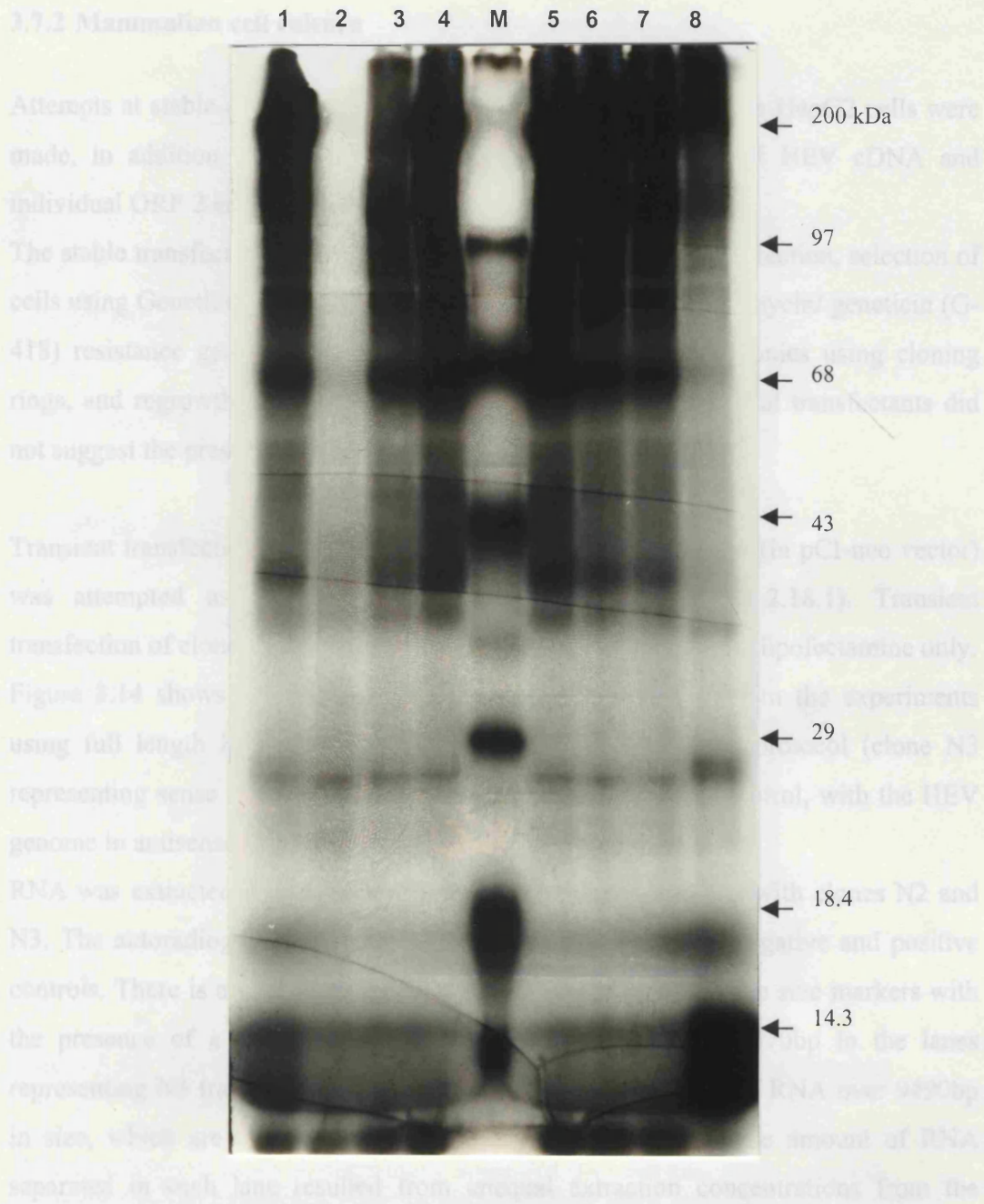
Figure 3.13: Expression *in-vitro* of full length HEV clone showing evolution of proteins with time

The full length HEV clone GC5 in the pCRII vector was expressed *in-vitro* using the Promega TnT system (SP6 polymerase version, method 2.17.6). The reaction used 1µg of plasmid in a 60µl reaction, with 5µl sampled at sequential times, and placed at -20°C for subsequent protein separation by SDS-PAGE using a gradient stacking gel. Autoradiography was for 5 days at -70°C.

The lanes are as follows:

M. ^{14}C molecular weight markers (Gibco BRL, UK), sizes as indicated by arrows
1. Positive control (previous reaction using GC5 TA full length HEV clone)
2. At time 15 mins **3.** At time 30 mins **4.** At time 45 mins **5.** At time 60 mins
6. At time 75 mins **7.** At time 90 mins **8.** Overnight incubation including 0.1mM puromycin

3.7.2 Mammalian cell culture

Attempts at stable expression of full length cloned HEV cDNA in HepG2 cells were made, in addition to transient expression of full length cloned HEV cDNA and individual ORF 2 and 3 in COS 7 cells.

The stable transfection protocol involved calcium phosphate transfection, selection of cells using Geneticin (since the pCI-neo vector contained the neomycin// geneticin (G-418) resistance gene, see appendix 5), isolation of cellular colonies using cloning rings, and regrowth under selective pressure. Analysis of potential transfectants did not suggest the presence of cloned HEV.

Transient transfection of full length cloned HEV genomic cDNA (in pCI-neo vector) was attempted using lipofectin and lipofectamine (method 2.18.1). Transient transfection of cloned ORF2 and 3 sequences was attempted using lipofectamine only. Figure 3.14 shows northern blot analysis of RNA extracted from the experiments using full length HEV cDNA transfection with the lipofectin protocol (clone N3 representing sense orientation and N2 serving as the negative control, with the HEV genome in antisense orientation).

RNA was extracted from three separate sets of cells transfected with clones N2 and N3. The autoradiograph shows a clear difference between the negative and positive controls. There is a diffuse signal between the 1350bp and 4400bp size markers with the presence of a more distinct band between 1350bp and 2370bp in the lanes representing N3 transfection. These lanes also show two bands of RNA over 9490bp in size, which are not visible in the control lanes. The variable amount of RNA separated in each lane resulted from unequal extraction concentrations from the different sets of transfected cells.
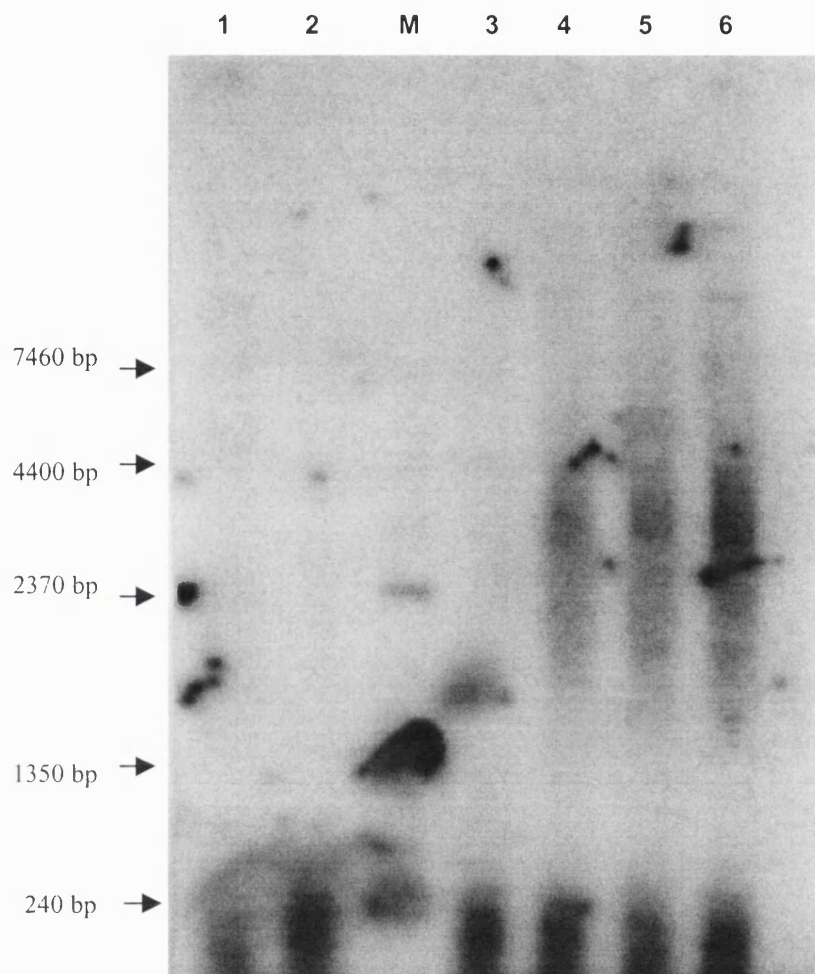
Figure 3.14: Northern blot of total RNA from COS 7 cells following transient transfection of a full length HEV clone

The autoradiograph shows Northern blot analysis of COS 7 cell total RNA, extracted following transient transfection using the full length HEV clone (method 2.18.1, lipofected cells). Autoradiography was for 6 days at -70°C. Blot processing was done as detailed in method 2.20.5. Two ^{32}P labelled DNA probes were used- 30ng lamda bacteriophage and 25ng of 1.4 Kb GC5 clone fragment spanning the terminal 3' end of HEV.

The lanes represent:
M. Poly-A tailed molecular weight markers (Gibco BRL, UK, 750ng).
1. 0.8µg RNA of clone N2 (antisense negative control) transfected cells, dish 1
2. 1.7µg RNA of clone N2 (antisense negative control) transfected cells, dish 2
3. 1.3µg RNA of clone N2 (antisense negative control) transfected cells, dish 3
4. 1.6µg RNA of clone N3 (sense orientation) transfected cells, dish 1
5. 1.1µg RNA of clone N3 (sense orientation) transfected cells, dish 2
6. 1.4µg RNA of clone N3 (sense orientation) transfected cells, dish 3

Figure 3.15 shows northern blot analysis of the total cellular RNA extracted following transient transfection, using the lipofectamine method, of cloned full length HEV cDNA, and cloned HEV ORF 2 and 3.

The negative control (lane 1, antisense orientation cloned HEV ORF 3) lane shows no signal. Lane 4, representing transfection of the cloned ORF 3 in sense orientation also shows no bands as the probe used in this experiment was derived from a region not spanning ORF 3. For the full length HEV genome transfection, there are two clear bands, one above 10 kb and one around 6 kb in size. The ORF 2 experiment also generated strong signals for RNAs of around 10 kb, 4 kb and 1.5 kb.
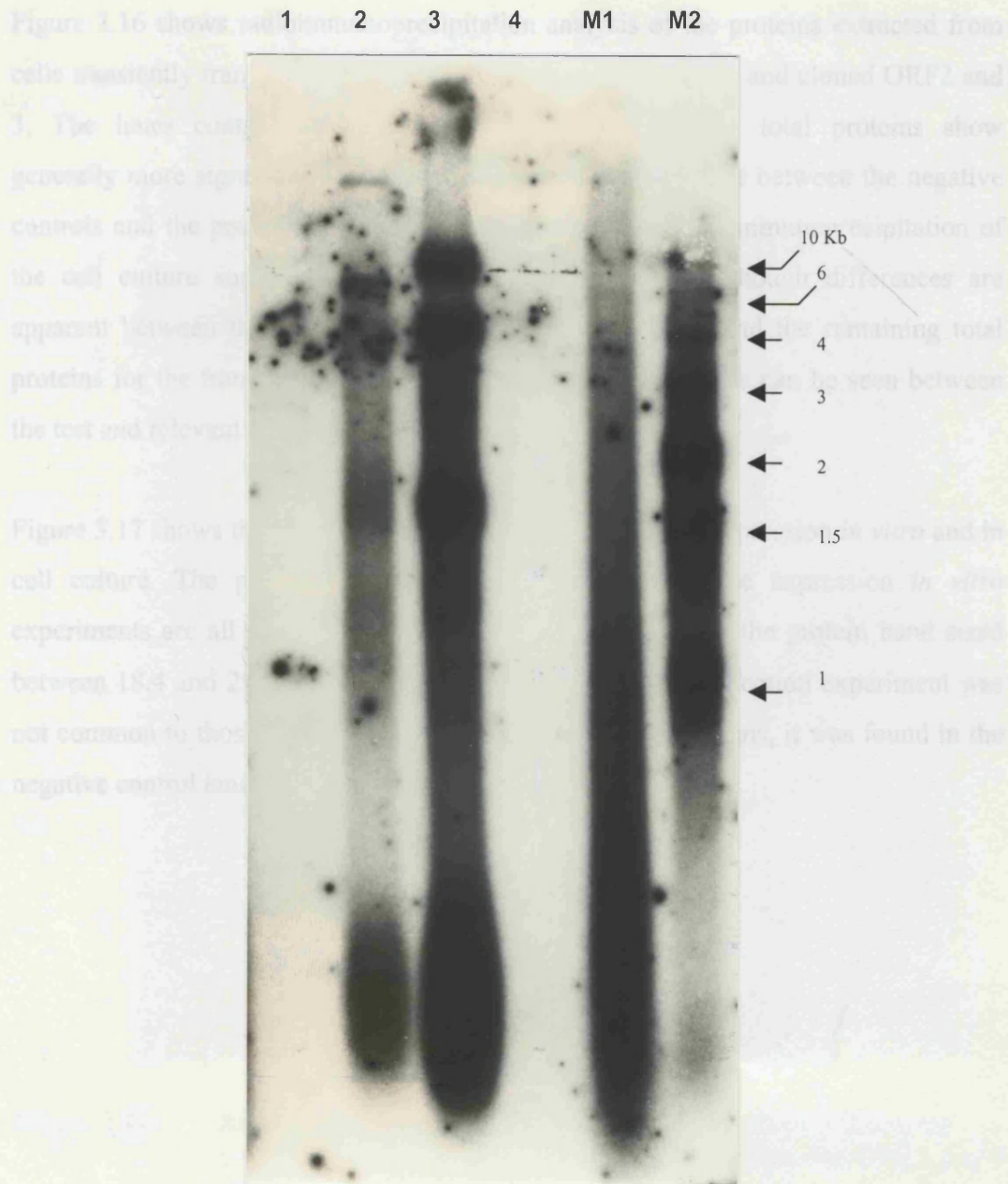
Figure 3.15: Northern blot analysis of total RNA from COS 7 cells following transient transfection of full length, ORF 2 and ORF 3 HEV clones

The autoradiograph shows Northern blot analysis of COS 7 cell total RNA, extracted following transient transfection with cloned full length genomic HEV cDNA, cloned ORF2 and cloned ORF3 (method 2.18.1, lipofectamine). Autoradiography was for 29 days at -70°C. Blot processing was done as detailed in method 2.20.5. A single ^{32}P labelled DNA probe (GC5 1.4 kb fragment positioned at terminal 3' end HEV) was used.

The lanes represent:

1. 700 ng RNA negative control (antisense cloned ORF3)
2. 500ng RNA from full length HEV clone transfection
3. 300ng RNA from ORF2 transfection   4. 700ng RNA from ORF3 transfection
M1. Molecular weight marker, ^{32}P end-labelled RNA M2. unlabelled RNA; sizes as indicated

Figure 3.16 shows radioimmunoprecipitation analysis of the proteins extracted from cells transiently transfected with cloned genomic HEV cDNA, and cloned ORF2 and 3. The lanes containing the radiolabelled (unprecipitated) total proteins show generally more signal as expected. There is no clear difference between the negative controls and the proteins from the transfection experiments. Immunoprecipitation of the cell culture supernatant yielded very little protein. Although differences are apparent between the immunoprecipitated cellular proteins and the remaining total proteins for the transfections of the same clone, no differences can be seen between the test and relevant control experiment.

Figure 3.17 shows the western blot of proteins generated by expression *in vitro* and in cell culture. The pattern of the proteins detected from the expression *in vitro* experiments are all similar to the negative control. Although the protein band sized between 18.4 and 29 kDa detected from the cell culture transfection experiment was not common to those found in the *in-vitro* expression extractions, it was found in the negative control lane and is therefore unlikely to be significant.
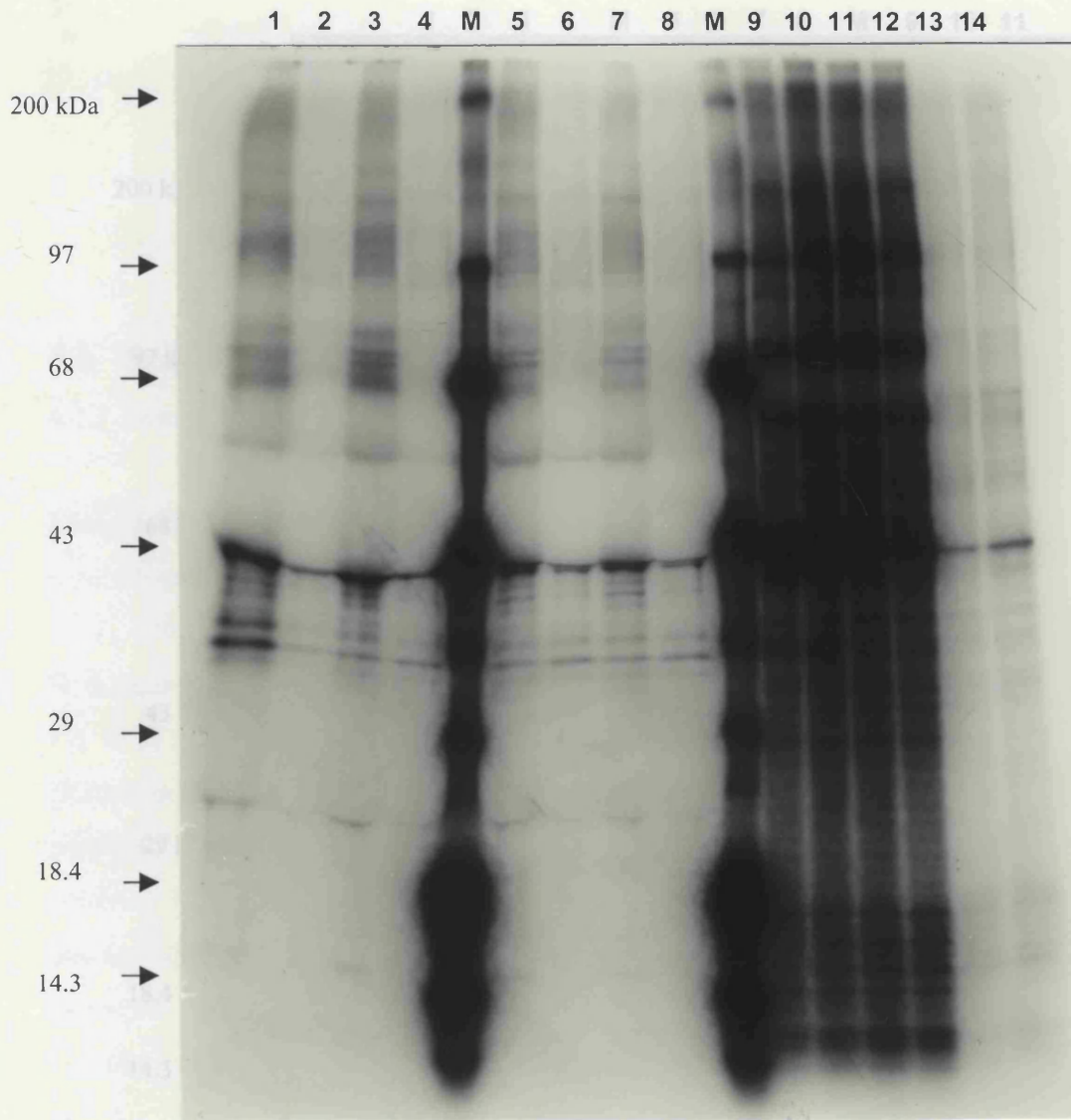
Figure 3.16: Radioimmunoprecipitation of COS 7 cellular proteins following transient transfection with full length HEV clones and ORF 2 and 3 clones

Autoradiograph of SDS-PAGE of radioimmunoprecipitated proteins (method 2.19.4) following transient transfection of cloned HEV sequences.
Lanes 1-8 immunoprecipitated proteins, 5 out of 40μl loaded
Lanes 9-14 remaining total proteins, 5 out of 180μl loaded

1. Cellular proteins full length HEV transfection **2.** Supernatant proteins full length HEV
3. Cellular proteins ORF 2 transfection **4.** Supernatant proteins ORF 2 transfection
5. Cellular proteins ORF 3 transfection **6.** Supernatant proteins ORF 3 transfection
7. Negative control cellular proteins (pCI-neo vector with antisense insert)
8. Negative control supernatant **9.** Cellular proteins full length HEV transfection
10. Cellular proteins ORF 2 transfection **11.** Cellular proteins ORF 3 transfection
12. Negative control cellular proteins (pCI-neo vector with antisense insert)
13. Supernatant proteins full length HEV **14.** Negative control supernatant

Figure 3.17: Western blot analysis of proteins derived from *in-vitro* translation and cell culture transient transfection of cloned HEV sequences

The western blot was done as detailed in method 2.19.3.
The lanes represent:
1-7: proteins extracted from *in-vitro* translation experiments using in pCI-neo vector with T7 promoter (method 2.17.4). 8μl of 25μl loaded.
8-11: proteins extracted from cell culture transient transfection experiments. Cells from lipofectamine protocol. 1x 35mm dish of cells resuspended in 200μl 1x sample buffer, boiled 3 minutes, 15μl loaded.

1. Expression of full length HEV.   **2.** Expression of full length HEV with microsomes
3. Expression of ORF 2.   **4.** Expression of ORF 2 with microsomes
5. Expression of ORF 3.   **6.** Expression of ORF 3 with microsomes
7. Negative control (antisense ORF 3 clone)
8. Transfection of full length HEV. **9.** Transfection of ORF 2.
10. Transfection of ORF 3.   **11.** Negative control (COS7 pCI-neo vector only)
M. Molecular weight markers, sizes indicated by arrows (Gibco BRL, UK)

146

CHAPTER FOUR

DISCUSSION

4.1 Sequence generation and analysis

4.1.1 Source of viral nucleic acid

Many early studies of HEV used viral material taken from monkeys, often passaged to obtain high titres prior to characterisation (Reyes *et al.,*1990; Tsarev *et al.,* 1992; Fry *et al.,* 1992; Aye *et al.,* 1992; Huang *et al.,* 1992; Aye *et al.,* 1993; Yin *et al.,* 1994). Some later studies used human stool as the source of HEV RNA (Huang *et al.,* 1992; Fry *et al.,* 1992; Bi *et al.,* 1993-affinity capture IgM; Wang *et al.,* 2000), or serum (Kabrane-Lazizi *et al.,* 2001- where some sequences were also derived form virus passaged in monkeys), but no others were found where liver was used as the source material. No consistent changes in HEV nucleotide sequence have been attributed to passage in monkeys, however, a variant HEV has resulted from a single passage in a macaque (van Cuyck-Gandre *et al.,* 1998). In that experiment, monkeys were inoculated intravenously with clarified stool taken from a case of hepatitis E in Chad, and the HEV nucleic acid extracted from the animal bile was sequenced and compared to the original inoculum and reference strain sequences. Variation in the peptide sequence of part of ORF 2 occurred after a single passage, with the most divergent sequence found in the animal showing the most severe hepatitis. It is uncertain whether this represents an artefact of amplification of a subpopulation or evolution of the virus *in-vivo*. Therefore, infected liver was chosen as the direct source of viral nucleic acid in order to eliminate possible sequence anomalies introduced by passage in monkeys.

The observational link between acute liver failure and HEV infection in the later stages of pregnancy suggests that host factors are probably key to the outcome in that clinical setting. Therefore, the use of viral nucleic acid taken from a male patient with acute liver failure due to HEV, eliminated any factors specific to pregnancy. No other

HEV nucleotide sequences deposited in Genbank are assigned to cases of acute liver failure.

4.1.2 Validity of nucleotide sequence data

Nucleic acid amplification by the polymerase chain reaction can lead to the generation of nucleotide sequence artefacts, the rate of which is dependent upon the type of thermostable DNA polymerase used. In order to reduce the impact of these possible sequence errors, a consensus HEV genomic sequence was constructed using data from several different methods, and from both strands of DNA (see figure 3.5). In particular, HEV contains several regions of GC rich sequence that can be problematic for nucleotide sequencing and amplification. One PCR reaction led to an HEV specific product containing a deletion (set 8, see section 3.1). Such a deletion would affect the ORF 2 signal sequence and therefore have a negative effect on replication competence. Comparison of the primer sequence to the final consensus sequence for our strain indicated that the primers were likely to be suboptimal. This area was later amplified and sequenced from full length product and partially from primer set 4 PCR products. It was unclear whether this finding represented a natural variant sub-population or an artefact of amplification. A similar in-frame deletion was reported earlier by a different group, comprising 246 bp in ORF 2/3 (between 5187 and 5433) of an Indian strain of HEV (Ray *et al.*, 1992).

Several of the short amplicon PCRs failed, primarily as a result of using published data in the design, resulting in mismatches. Other PCR reactions generated products of the correct size as well as additional products. In the case of the genomic PCR, amplicons of around 3.7 kb and 2.0 kb were produced, however, the sequence of these was not defined and the study strain did not contain any regions likely to result in mispriming. Although selection of the correctly sized amplicons for cloning may have biased the results, the overall sequence data was confirmed from several full length HEV clones.

4.1.3 Nucleotide and amino acid sequence comparisons

The HEV study strain genome conformed to the organisation of previously described prototype strains from Burma, Pakistan and Mexico, and contained all of the motifs expected (see 3.6) and there was broad conservation of the immunoreactive epitopes identified by various groups.

The lengths of the individual ORFs and terminal non-coding regions were similar to those given for the Burmese prototype strain (Tam *et al.,* 1991), with the exception of the 5' NCR which was 1 nucleotide shorter in our study strain (26 versus 27 nucleotides, see results table 3.1). The overall closest matches in the databases were with members of genotype 1, within which there was a high degree of conservation at both the nucleotide and amino acid level. Comparisons of both the 5' and 3' NCRs for published sequences were made, since these regions might be considered important in viral replication regulation.

Figure 3.6 shows the pileup of 5' NCRs, clearly showing the that the study strain is unique. For genotype 1 sequences only the study strain and the Chinese isolate C2 lacked an initial adenosine, and only the study isolate had a cytosine at position 11. The 5' NCR appears otherwise well conserved for the genotype 1 members, seven of which were identical, but no other isolates from cases of acute liver failure have been described for comparison. The importance of the 5' NCR in viral replication competence has not been determined, and for some strains this area has not been identified at all (US1- AF060668, Schlauder *et al.,* 1998) or it is unusually long (US2-AF060669). This might suggest that the 5' NCR is not critical in replication, however, it is unclear whether this simply reflects an inability to detect that region or whether it is truly absent. Until recently, no 5' NCR had been described for the Pakistani prototype strain (P1- M80581) and that for the Mexican strain (M1- M74506) was thought to be truncated, whereas new data has shown them to be similar to the Burmese prototype strain (B1- M73218) (Zhang *et al.,* 2001).

The 3' NCR for the study strain was not unique, as it was identical to one other strain (I2- X99441 von Brunn, unpublished) that had not been associated with acute liver failure. This area showed considerable divergence across genotypes and its role in viral replication is still to be established.

In order to test the importance of these terminal genomic regions in pathogenicity, a suitable animal model or cell culture system would be required, allowing infection using strains engineered to contain the study strain 5' and 3' sequences on a background of viral genome of known pathogenicity.

Table 3.2 depicts the similarities at the nucleotide level for different full length genotype representatives, and highlights the basis of the genotype groupings. The study isolate has between 92 and 94% similarity with members of the same genotype (1), whereas this was redued at 74.4 to 75.7% for genotypes 2, 3 and 4. Within genotype 1, other members showed much closer similarities, for example, the Chinese isolate C2 and the Pakistan isolate P1 showed 98.8% identity.

As expected, the amino acid identities (tables 3.3, 3.4, and 3.5) were higher than those for the nucleotide values due to silent third base in codon changes; ORF 1 of the study strain showed between 95.5 and 98.4% similarity to other members of genotype 1, for ORF2 the values were 97.1 to 98.8% and for ORF3 they were 94.3 to 98.4%. The highest identities for ORF 2 were maintained for comparison against other genotypes, this relative conservation probably reflecting the putative role of ORF 2 as a capsid protein. Reduced similarities for ORF 1 across genotypes is likely to be a function of the hypervariable region.

Different strains showed the closest homology with the study strain depending on the ORF compared. This was confirmed by analysis of the study strain by another group, showing that the ORF 2 region had highest homology with a Burmese strain, whereas in the ORF 1 region the closest match was with a Chinese strain, the relevant Burmese strain showing significantly lower identity (Tsarev *et al.,* 1999).

With the widespread use of sensitive PCR methods for the detection of HEV, new data are appearing all the time, and such comparisons are likely to become more complex.

4.1.4 Phylogenetic analysis

Many authors have constructed phylogenetic trees for use in assigning genotypes to published HEV strains. Where this has been done with full length HEV sequences, four genotypes have been proposed. In common with other viruses, HEV isolates contain regions of highly conserved sequence across genotypes in addition to regions

which are more variable. For this reason, genotype definitions based upon comparisons of short nucleotide sequences across a single genomic region may be misleading. A strain of HEV derived from a traveller who had been to Thailand was found by nucleotide identity to be more closely related to or more distant from a new Chinese isolate (genotype 4) than other Asian strains according to the region used (Kabrane-Lazizi *et al.,* 2001). However, such analyses may be necessary when attempting to include as many strains as possible in a phylogenetic tree. In addition, data taken from Genbank or other sources often does not contain precise information about the geographical source of the virus, making comments regarding clustering difficult.

Phylogenetic analysis has been used to suggest the division into subtypes (Wang *et al.*, 1999), for example, for genotype 1, subtype 1a includes B1 M73218, B2 D10330, I2 X99441; subtype 1b includes C1-6, D11092, L2554, M94177, D11093, L08816, L25595, and P1, M80581; subtype 1c, I1 X98292. Such analysis has been used to show a shift in subgenotype in epidemic strains over a time period in the same area (Arankalle *et al.*, 1999). These divisions have not been standardised, with different authors presenting different findings (Arankalle *et al.*, 1999; Wang *et al.,* 1999; Tsarev *et al.*, 1999; Schlauder and Mushahwar, 2001). Appendix 11 shows a table of proposed HEV genotype classifications.

The study strain has been grouped within genotype 1 by several authors (Arankalle *et al.*, 1999; Wang *et al.*, 1999; Tsarev *et al.*, 1999), however often showing the highest degree of divergence within members of the same group.

The analysis in figure 3.8 confirms the genotype groupings previously described, presented as a rooted phylogenetic tree. The Pakistan, Chinese and Indian isolates are clustered, the two US types are clustered and the Chinese T1 and Mexican isolates form single branches (expected for the Mexican isolate as this was chosen as the outgroup). The groupings appear very robust, with bootstrapping percentages of a hundred replicates not less than 88%. In addition, the tree reflects the percentage nucleotide similarities shown in table 3.2, as can be seen by the tightest cluster of sequences C1-4 which are all over 96% identical. Analysis of full length genotype 1 sequences by another author (Erker *et al.*, 1999) resulted in similar findings to those in figure 3.8. The study strain appeared as a single branch within the genotype 1 cluster in both trees in that analysis, despite the use of different programs to derive the relationships. The significance of this is unclear, although it could relate to differences

between sporadic (I1) and epidemic (others) strains, or simply be a geographical phenomenon. HEV subtypes have been shown to change in the same geographical area over time (Arankalle et al., 1999), and it is likely this occurs between areas also (sequence I2 was derived form an infection in Madras, the specific origin of I3 is not given and infection with the study isolate may have occurred in Bombay, Delhi, Bangalore or Goa).

Figure 3.9 shows the phylogenetic tree of genotype 1 representatives based on a relatively conserved region of ORF 2, partly defined by primers described in Wang et al., 1999. This area was chosen to allow comparison of the maximum of data. Prior to alignment, all non-genotype 1 sequences were removed. The tree shows that the differences between isolates are small, however there is some geographical clustering of Burmese strains (genbank accessions D10330, D90274), as well as Egyptian strains. The Chinese isolates with genbank accession numbers L25595 and L25547 appear very closely related and nucleotide analysis suggests that they may in fact be identical. In agreement with other authors' findings (Arankalle et al., 1999; Erker et al., 1999), the sudy strain I1 has grouped independently. The Moroccan sequence is branched alone as this was chosen as the forced outgroup. HEV appears to be genetically stable during epidemics and across time in the same areas, with only minor changes in subgenotype found. When different genotypes are found co-circulating, this is likely to be the result of recent importation of a new type.

The overall similarity of the genotype 1 representatives may have led to the generally low bootstrap values generated. Under such circumstances, several clusters of sequences may be equally valid. Comparison of an alternative area of HEV sequence may have helped define the clusters presented in figure 3.9, however, no other genomic regions could be found that would include all the isolates necessary.

Consideration was given to construction of a phylogenetic tree relating the amino acid sequence of the ORF 2 region selected, however the conserved nature of the area meant that the majority of strains were identical, rendering the analysis unhelpful.

Data alignment was required prior to phylogenetic analysis and this led to removal of some sequence at the extreme 5' and 3' ends of some isolates, which may have introduced bias into the results. Despite this, the findings agree broadly with those of other authors using different techniques.

Genotyping may show the relationship between human and non-human HEV isolates, and reflect the geographical evolution over time of HEV. Since no single genotype of HEV has been shown to be linked with a particular outcome of infection, and since there are currently no effective antiviral agents for HEV hepatitis, it is unclear what clinical value genotyping can provide. For this reason, a simpler method of classification would seem appropriate, for example, scheme 1 listed in appendix 11 and discussed in Schlauder and Mushahwar, 2001. This scheme is used in the data presented in this thesis.

4.2 Replication strategy of HEV

There is still no efficient cell culture method for HEV, and this has prevented clear analysis of the replication strategy of HEV. Despite this, there are data from various studies describing the nature of the proteins expressed by full length genomes or from individual ORFs in translation systems. Our work also attempted to define the sizes of HEV specific proteins and RNAs generated *in-vitro* or by mammalian cell culture.

4.2.1 HEV proteins

The expression *in-vitro* of full length genomic HEV sequence (figures 3.10 and 3.11) resulted in several proteins, one of which was close to 200 kDa in weight. This may represent the entire ORF 1 product, since ORF 1 seems to encode a protein of around 186 kDa when expressed alone (Ansari *et al.,* 2000; Ropp *et al.,* 2000), corresponding to the size predicted from nucleotide translation. Further support for this as an HEV specific protein can be seen in figure 3.10, which shows the translation of clone A3 (lane 1) producing a smaller protein between 68 and 97 kDa, corresponding to the size predicted (83 kDa) by nucleotide translation up to the premature stop codon at position 2297 in that clone; the larger protein is absent. Although ORF 1 encodes a motif similar to a papain-like cysteine protease, expression of ORF 1 *in-vitro* using coupled transcription/ translation has not been shown to yield smaller proteins by proteolytic processing, even in the presence of co-factors such as divalent cations or microsomes (Ropp *et al.,* 2000). Two experiments were done to test for co-translational or post-translational modification of the HEV proteins produced by *in-*

vitro expression of the study strain HEV. The results of expression in the presence of protease inhibitors (figure 3.12) were difficult to interpret due to the poor reaction efficiency, however comparison between the test and control lanes did not indicate protease activity. Sampling of an *in-vitro* expression reaction at incremental times (figure 3.13) did not show gradual accumulation of proteins of increasing size and did not demonstrate appearance of smaller products over time, that would have been indicative of cleavage. Very little translation had taken place at time 15 minutes, however, at 30 minutes the putative full length ORF 1 protein was present, and at 45 minutes the reaction appeared complete. Overnight incubation in the absence of translation resulted in non-specific degradation of the proteins. In comparison, other workers demonstrated appearance of the 185 kDa protein at 40 minutes (Ropp *et al.*, 2000). In order to gain further information, additional reaction sampling would be required between 15 and 30 minutes. The proteins generated *in-vitro* outside the size ranges predicted for the individual ORFs are either vector related, or may be the result of internal initiation or premature termination of translation.

Expression of ORF 1 in cell culture systems has given additional information, however, sometimes with conflicting findings. Despite the presence of the motif similar to a papain-like cysteine protease, such pocessing has not been observed following expression of ORF 1 alone in *E. coli* or HepG2 cells (Ansari *et al.*, 2000). However, extended expression (24-36 hours) of ORF 1 as part of recombinant vaccinia virus system in cell culture (HeLa, vero, HepG2 and human Chang liver) did result in 107 and 78kDa products (Ropp *et al.*, 2000). In that study, mutagenesis of the Cys 483 amino acid, predicted to be part of the catalytic target of papain-like cysteine proteases, did not abolish cleavage. Given that the other predicted target His 590, was not conserved across HEVs, it is possible that HEV is not processed by the papain-like cysteine protease or it uses a novel virally encoded protease at a single site only. Immunoprecipitation of products labelled at either the C or N-termini with peptide tags (FLAG peptide and anti-Tag antibody) showed that many smaller proteins formed were premature termination products. Similar immunoprecipitation experiments using proteins generated in cell culture showed the 107 kDa protein to be N-terminal and the 78 kDa, C-terminal. Where a full length cDNA clone of HEV has been expressed in HepG2 cells (Panda *et al.*, 2000), immunoprecipitation revealed the presence of proteins thought to represent the methyltransferase (35kDa), helicase (38kDa), and polymerase (36kDa). The polymerase appeared to be an early product,

detectable at 12, 24, and 36, but not 72 and 96, hours post-transfection. The helicase and methyltransferase were found at 72 and 96 hours. Evidence of cellular replication came from detection of HEV negative strand RNA, which was present in lower concentration than the positive (genomic) strand. Therefore, ORF 1 seems to undergo processing in the context of full genome expression, but not when expressed alone *in-vitro* or in HepG2 cells. Such processing occurred despite the addition of 12 nucleotides at the 5' end, one at the 3' end, and the absence of a cap structure in the full length transcript.

Transient transfection of COS7 cells with the study strain HEV did not lead to production of HEV specific protein detectable by immunoprecipitation (figure 3.16). This could be accounted for by either a low efficiency for the transfection or a suboptimal detection method, or a combination of the two. Alternatively, the cells may have been harvested too late, at a time when many cells had died, leading to RNA degradation. Since it was unclear whether HEV specific RNA was present in transfected cells (figures 3.14 and 3.15), this may suggest failed transfection. Transfection protocols using cationic lipids often require optimisation of the relative proportions of transfection agent and DNA. As only one experiment was possible due to time constraints, the amount of reagent used was in the middle of the range suggested by the manufacturer. Altering the concentration may have improved the results. The immunoprecipitation relied on antibody derived from the study patient's serum taken during the acute phase illness, and although there may have been relatively more HEV IgM present, HEV IgG was clearly detectable (Fagan *et al.*, 1994). The above findings show that any viral protease activity has not yet been characterised.

Many *in-vitro* and cell culture expression experiments have been done to determine the nature of the proteins encoded by ORF 2 and ORF 3, partly resulting from their potential use in antibody detection assays and as vaccine candidates. The putative capsid protein is thought to be encoded by ORF 2, which has a predicted protein weight of around 72 kDa. Expression *in-vitro* of ORF 2 has yielded products of between 72 and 74 kDa (Jameel *et al.*, 1996). Expression *in-vitro* of the study strain ORF 2 and 3 independently was originally intended to to be done directly from PCR products, however a low concentration of amplicon meant plasmid cloning was required. Figure 3.11 shows the expression of cloned ORF 2 resulting in a major

product of around 70 kDa and this was also produced in reduced amount by the full-length HEV clone (figures 3.10 and 3.11). The proteins sized between 68 kDa and 14 kDa most likely are accounted for by a mixture of vector translation and internal initiation or premature termination of HEV sequence translation. Only a small proportion of non-vector related proteins are shared between the full-length genomic translation and that for ORF 2 alone. ORF 2 contains three potential N-linked glycosylation sites (represented by Asn-X-Thr/Ser) at 137, 310, and 561. The addition of microsomes did not result in a change in gel mobility for the 70 kDa ORF 2 product, however, they did markedly reduce the reaction efficiency.

Cell culture expression of ORF 2 by other groups has produced a variety of products which could be explained by post-translational modification effects. Pulse chase expression studies of ORF 2 alone in COS-1 cells using an SV40 based vector appeared to produce a precursor protein of 82 kDa, the non-glycosylated ORF 2 product of 74 kDa and a glycosylated form at 88 kDa (Jameel *et al.*, 1996). The different forms of ORF 2 protein run anomolously by SDS-PAGE, with the primary sequence product of predicted 72 kDa running at 82 kDa and the signal cleaved form running at 74 kDa. This effect may result from the presence of glycosylation and the proline rich sequences. In addition, pORF2 may form dimers that can be dissociated by heating at 100°C for 2 minutes prior to electrophoresis. Gel sizing may therefore be unreliable.

The ORF 3 protein is thought to be 13.5 kDa, without glycosylation (Jameel *et al.*, 1996; Panda *et al.*, 2000). The role of pORF3 in viral replication is unclear as it is uncertain whether it forms part of the virion, and pORF2 alone can form virus like particles. Interaction with a cellular protein has been suggested, since immunoprecipitation of ORF 3 protein led to co-precipitation of an 18 kDa protein if expression was in COS-1 or Hu7 hepatoma cells, however this was not observed with HepG2 cells (Jameel *et al.*, 1996). The role of this cellular protein is uncertain. The differences between the cell lines could result from the different stages of differentiation. Figure 3.11 shows the expression *in-vitro* of the cloned ORF 3 study strain sequence, resulting in two clear non-vector proteins, one between 14 and 18 kDa and the other approximately 14 kDa. The 14 kDa protein probably represents the ORF 3 product.

Although the size of some of the proteins produced by expression *in-vitro* suggests they are HEV related according to sequence prediction and the findings of other

groups, further confirmation was required. Immunoprecipitation of cell culture derived proteins was negative and it was unclear whether this was the result of failed transfection or suboptimal detection. Western blotting (figure 3.17) of cell culture proteins and proteins generated by expression *in-vitro* did not demonstrate HEV proteins either. Since the latter reactions contained proteins with sizes predicted for the HEV ORFs, and proteins of similar sizes had transferred during blotting, it is likely that the detection system may not have been adequate. The serum used did contain HEV IgG (Fagan *et al.*, 1994) but was only available in limited amount, and repeated experiments using increased antibody concentration or altered blot washing conditions were not possible. Due to the rarity of imported HEV infection in the UK, and insufficient time to establish a source from a country where HEV was endemic, further anti-HEV sera were not easily obtained.

4.2.2   Detection of HEV RNAs

There have been reports of the detection of HEV subgenomic RNAs (Tam *et al.*, 1991; Yarbough *et al.*, 1991). In addition to the full length HEV genome of 7.5 kb, RNAs of around 2 kb and 3.7 kb were detected by northern blotting of liver RNA taken from experimentally infected cynos. Due to the nature of the probe used, these subgenomic RNAs were thought to be co-terminal and located at the 3' end of the genome, however, the ends of these RNAs were not defined and there have been no subsequent confirmatory reports. We attempted to detect HEV RNA in the total RNA extracted from the explanted liver, using probes designed to cover the 3' end of the genome. The northern blot in figure 3.4 shows the presence of RNA bands at around 5 kb and 2 kb, corresponding to the size of ribosomal RNA. The lack of detection of full length genome and the possible subgenomics may be the result of massive necrosis of the explanted liver and RNA degradation. Northern blotting of poly-A selected RNA may have led to improved specificity, however, the source material was limited and this could not be done. Since we could not exclude the 5 kb and 2 kb RNAs as subgenomic HEV, 5' RACE experiments were done to try to locate the position of potential subgenomics (see below).

In order to check for HEV transcription following COS7 cell transient transfection, total RNA from transfectants was harvested and northern blotted (figure 3.14). Although the amounts or RNA loaded per lane were not identical, there is a clear

difference between the negative control cells and the test cells. Transfection of the study strain HEV plasmid has resulted in the presence of several RNA bands not seen in the controls. There is a broad band between 2 and 5 kb, as well as two bands above 7.4 kb. Therefore, no bands were detected corresponding to full length HEV genome. The larger RNAs may represent HEV sequence and vector combined, as the total transcript would be 12.65 kb, and the broad band may contain degraded RNA and possibly subgenomic message. As these RNAs were not seen in the negative controls, that were transfected with vector containing HEV genome in antisense orientation, they were not thought to be cellular in origin. Further northern blotting was done for RNA extracted from cells transfected with plasmids containing full length HEV and ORF 2 and 3 alone (figure 3.15). Transfection was done using an improved method (lipofectamine solution, method 2.18.1). Transfection of cloned full length genome and ORF 2 generated RNAs not detected in the control. For the full length genome transfection, an RNA of around the expected size of 7 kb is apparent, in addition to a larger band of over 10 kb. The ORF 2 transfection produced RNAs of around 10 kb, between 4 and 6 kb, and between 1.5 and 2 kb (also generated by the full length genome transfection). Since the expected size of ORF 2 is 2 kb, the nature of the detected RNAs is uncertain. It should be noted that in the first paper proposing the presence of subgenomic message (Tam et al., 1991), northern blotting of the liver RNA did detect RNAs above 7.5 kb, although these are not discussed.

The 5' RACE experiments successfully identified the 5' end of the genomic HEV RNA and this technique was used to attempt to define the 5' location of potential subgenomic HEV RNA. Multiple primers were chosen to amplify sections of HEV cDNA of around several hundred bases, spanning the regions where the subgenomic RNAs were thought to have the 5' end (table 2.3). Repeated attempts failed to clearly show the presence of HEV subgenomic RNA, concurring with the results of northern blotting. Once again, this may have been affected by the quality and amount of HEV RNA remaining in the explanted liver, however the presence of sufficient full length HEV genome to be amplified by PCR makes this less likely.

The information generated from various HEV expression experiments has still not led to a defined mechanism by which HEV replicates, particularly in relation to expression of ORF 2 and 3. Disparate groups of plant and animal viruses with positive

strand RNA can share similar replication strategies, although different host cell factors are likely to be critical (Pogue *et al.*, 1994). Molecular analysis of the HEV genome (Koonin *et al.*, 1992), and experimental analysis of the function of the HEV methyltransferase (Magden *et al.*, 2001) has indicated that HEV may belong to an alphavirus-like supergroup. The replication strategy of brome mosaic virus (BMV), a member of the alphavirus superfamily, has been described. BMV contains an internal control region (ICR) at the 5' end that functions within the 5' stem loop structure essential for genome replication. The motif for the BMV ICR is GGTTCGANTCC, which resembles the part of the 5' cloverleaf stem loop structure for poliovirus, also necessary for infectivity. Similar motifs could not be identified in the HEV study strain, and its role in replication is unclear. HEV antisense genome has been detected in the livers of experimentally infected monkeys (Nanda et al., 1994) but the existence of subgenomic RNAs and their involvement in expression of ORF 2 and 3 is contentious. There are only two reports of such subgenomics, relating to northern blotting of infected monkey liver RNA (Tam *et al.*, 1991; Yarbough *et al.*, 1991), and their role in expression of ORF 2 and 3 has not been confirmed. Therefore, the overall replication strategy of HEV remains unproven.

4.3 Conclusion

This work includes the first description of an HEV genome nucleotide sequence derived directly from nucleic acid extracted from liver tissue explanted from a patient with acute HEV associated liver failure. The sequence conformed to the basic genomic stucture of previously described HEVs. Molecular analysis of the study strain HEV failed to demonstrate any clear areas that might be associated with increased pathogenicity. The 5' NCR was unique amongst published sequences, however the role of this region in replication is yet to be established. Translation *in-vitro* of the HEV cDNA generated proteins in the size range predicted by translation of the individual open reading frames.

Future work should include analysis of other strains of HEV taken from individuals with acute liver failure, where comparison might allow identification of putative areas responsible for altered pathogenicity. Sequence motifs suspected of playing a key role in pathogenicity could be inserted on a background HEV genome of standard pathogenicity. Wild type and recombinant viruses could then be used in cell culture or animal inoculation experiments to determine their effect *in-vivo*. Until such work is carried out, the current information indicates that the severe outcome of HEV infection seen in some cases is as likely to be a result of the host response as it is to relate to specific viral factors.

REFERENCES

Acharya S. K., Dasarathy S., Kumer,T. L., Sushma S., Prasanna K. S., Tandon A., Sreenivas V., Nijhawan S., Panda S. K., Nanda S. K., Irshad M., Joshi Y. K., Duttagupta S., Tandon R. K., Tandon B. N. (1996). Fulminant hepatitis in a tropical population: clinical course, cause, and early predictors of outcome. *Hepatology*. **23**(6): 1448-55.

Acharya S. K., Panda S. K., Saxena A., Datta Gupta S. (2000). Acute hepatic failure in India: A perspective from the East. *Journal of Gastroenterology and Hepatology*. **15**(5): 473-479.

Aggarwal R. and Naik S. R. (1992). Faecal excretion of hepatitis E virus. *Lancet* **340**(Sept 26): 787.

Aggarwal R. and Naik S. R. (1994). Hepatitis E: intrafamilial transmission versus waterborne spread. *Journal of Hepatology*. **21**: 718-723.

Aggarwal R., Kini D., Sofat S., Naik S. R., Krawczynski K. (2000). Duration of viraemia and faecal viral excretion in acute hepatitis E. *Lancet*. **356**: 1081-2.

Alter M. J. (1995). Epidemiology of hepatitis C in the West. *Seminars in Liver Disease*. **15**: 5-14.

Alter M. J., Gallagher M., Morris T. T., Moyer L. A., Meeks E. L., Krawczynski K., Kim J. P., Margolis H. S. (1997). Acute non-A-E hepatitis in the United States and the role of hepatitis G virus infection. *New England Journal of Medicine*. **336**: 741-746.

Altschul S. F., Madden T. L., Schafer A. A., Zhang J., Zhang Z., Miller W., Lipman D. J. (1997). Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic Acid Research*. **25**: 3389-3402.

Andre F. E., Hepburn A., D'Hondt E. (1990). Inactivated candidate vaccines for hepatitis A. *Progress in Medical Virology*. **17**: 72-95.

Andzhaparidze A. G., Balayan M. S., Savinow A. P., Braginsky D. M., Poleschuk V. F., Zamyatina N. A. (1986). Non-A, non-B hepatitis transmitted by the faecal-oral mode experimentally produced in monkey. *Voprosy Virusologii*. **31**: 73-81.

Ansari I. H., Nanda S. K., Durgapal H., Agrawal S., Mohanty S. K., Gupta D., Jameel S., and Panda S. K. (2000). Cloning, sequencing, and expression of the hepatitis E virus (HEV) nonstructural open reading frame 1 (ORF1). *Journal of Medical Virology*. **60**: 275-283.

Arankalle V. A., Ticehurst J., Sreenivasan M. A., Kapikian A. Z., Popper H, Pavri K. M., Purcell R. H. (1988). Aetiological association of a virus like particle with enterically transmitted non-A, non-B hepatitis. *Lancet*. **1**: 550-554.

Arankalle V. A., Chobe L. P., Jha J., Chadha M. S., Banerjee K., Favorov M. O., Kalinina T., Fields H. (1993). Aetiology of acute sporadic non-A, non-B viral hepatitis in India. *Journal of Medical Virology.* **40**: 121-125.

Arankalle V. A., Favorov M. O., Chadha M. S., Phule D. M., Banerjee K. (1993b). Rhesus monkeys infected with hepatitis E virus (HEV) from the former USSR are immune to subsequent challenge with an Indian strain of HEV. *Acta Virologica.* **37**: 515-518.

Arankalle V. A., Chadha M. S., Banerjee K., Srinivasan M. A., Chobe L. P. (1993c). Hepatitis E virus infection in pregnant rhesus monkeys. *Indian Journal of Medical Research.* **97**: 4-8.

Arankalle V. A., Chadha M. S., Tsarev S. A., Emerson S. U., Risbud A. R., Banerjee K., Purcell R. H. (1994). Seroepidemiology of water-borne hepatitis in India and evidence for a third enterically-transmitted agent. *Proceedings of the National Academy Sciences.* **91**: 3428-32.

Arankalle V. A., Goverdan M. K., Banerjee K. (1994b). Antibodies against hepatitis E virus in Old World monkeys. *Journal of Viral Hepatitis.* **1**(2): 125-129.

Arankalle V. A., Tsarev S. A., Chadha M. S., Alling D. W., Emerson S. U., Banerjee K., Purcell R. H. (1995a). Age-specific prevalence of antibodies to Hepatitis A and E viruses in Pune, India, 1982 and 1992. *Journal of Infectious Diseases.* **171**: 447-450.

Arankalle V. A., Chadha M. S., Chobe L. P., Nair R., Banerjee K. (1995b). Cross-challenge studies in rhesus monkeys employing different Indian isolates of hepatitis E virus. *Journal of Medical Virology.* **46**: 358-363.

Arankalle V. A., Chadha M. S., Dama B. M., Tsarev S. A., Purcell R. H., Banerjee K. (1998). Role of immune serum globulins in pregnant women during an epidemic of hepatitis E. *Journal of Viral Hepatitis.* **5**: 199-204.

Arankalle V. A., Chadha M. S., Chobe L. P. (1999). Long-term serological follow up and cross-challenge studies in rhesus monkeys experimentally infected with hepatitis E virus. *Journal of Hepatology.* **30**: 199-204.

Arankalle V. A., Paranjape S., Emerson S. U., Purcell R. H., Walimbe A. M. (1999b). Phylogenetic analysis of hepatitis E virus isolates from India (1976-1993). *Journal of General Virology.* **80**:1691-1700.

Arankalle V. A., and Chobe L. P. (1999). Hepatitis E virus: can it be transmitted parenterally? Journal of Viral Hepatitis. 6(2): 161-164.

Arankalle V. A., and Chobe L. P. (2000). Retrospective analysis of blood transfusion recipients: evidence for post-transfusion hepatitis E. *Vox Sang.* **79**: 72-4.

Arankalle V. A., Joshi M. V., Kulkarni A. M., Gandhe S. S., Chobe L. P., Rautmere S. S., Mishra A. C., Padbidri V. S. (2001). Prevalence of anti-hepatitis E virus antibodies in different Indian animal species. *Journal of Viral Hepatitis.* 8(3): 223-227.

Arif M., Qattan I., Al-Faleh F., Ramia S. (1994). Epidemiology of hepatitis E virus (HEV) infection in Saudi Arabia. *Annals Tropical Medicine Parasitology* 88: 163-168.

Arora N. K., Nanda S. K., Gulati S., Ansari I. H., Chawla M. K., Gupta S. D., Panda S. K. (1996). Acute viral hepatitis types E, A and B singly and in combination in acute liver failure in children in North India. *Journal of Medical Virology.* 48: 215-221.

Asher L.V. S., Innis B. L., Shrestha M. P., Ticehurst J., Baze W. B. (1990). Virus-like particles in the liver of a patient with fulminant hepatitis and antibody to hepatitis E virus. *Journal of Medical Virology.* 31: 229-233.

Aubry P., Niel L., Niyongabo T., Kerguelen S., Larouze B. (1997). Seroprevalence of hepatitis E virus in an adult urban population from Burundi. *American Journal of Tropical Medicine & Hygiene.* 57: 272-273.

Aye T. T., Uchida T., Ma X., Iida F., Shikata T., Zhuang H., Win K. M. (1992). Complete nucleotide sequence of a hepatitis E virus isolated from the Xinjiang epidemic (1986-1988) of China. *Nucleic Acid Research.* 20: 3512
(Genbank accession **D11092**)

Aye T. T., Uchida T., Ma X., Iida F., Shikata T., Ichikawa M., Rikihisa T., Win K. M. (1993). Sequence and gene structure of the hepatitis E virus isolated from Myanmar. *Virus Genes.* 7(1): 95-109.
(Genbank accession **D10330**)

Balayan M. S., Andzhaparidze A. G., Savinskaya S. S., Ketiladze E. S., Braginsky D. M., Savinov A. P., Polescuk V. F. (1983). Evidence for a virus in non-A/non-B hepatitis transmitted by the faecal oral route. *Intervirology.* 20: 23-31.

Balayan M. S., Usmanov R. K., Zamyatina N. A., Djumalieva D. I., Karas F. R. (1990). Brief report: experimental hepatitis E infection in domestic pigs. *Journal of Medical Virology.* 32: 58-59.

Balayan M. S., Zamyatina N. A., Ivannikova T. A., Khaustov V. I., Andzhaparidze A. G., Polescuk V. F. (1991). Laboratory models in the studies of enterically transmitted non-A, non-B hepatitis. In: Shikata T., Purcell R. H.. Uchida T. (eds). Viral hepatitis C, D, and E. Amsterdam: *Elsevier Scientific Publishers.* 213-220.

Balayan M. S., Fedorova O. E., Mikhailov M. I., Rytick P. G., Eremin V. F., Danilova T. I., Shevelev B. J., Gorbacheva E. C. Pankova G. Y. (1997). Antibody to hepatitis E virus in HIV-infected individuals and AIDS patients. *Journal of Viral Hepatitis.* **4**(4): 279-283.

Barzilai A., Schulman S., Karetnyi Y. V., Favorov M. O., Levin E., Mendelson E., Weiss P., Fields H. A., Varon D., Martinowitz U. (1995). Hepatitis E Virus Infection in Haemophiliacs. *Journal of Medical Virology.* **46**: 153-156.

Belabbes E. H., Bourguermouth A, Benatallah A, Illoul G. (1985). Epidemic non-A, non-B hepatitis in Algeria: strong evidence for its spreading by water. *Journal of Medical Virology.* **16**: 257-263.

Benster B. and Wood E. J. (1970). Immunoglobulin levels in normal pregnancy and pregnancy complicated by hypertension. *Journal of Obstetrics & Gynaecology of the British Commonwealth.* **77**(6): 518-22.

Berke T., and Matson D. O. (2000). Reclassification of the *Caliciviridae* into distinct genera and exclusion of the hepatitis E virus from the family on the basis of comparative phylogenetic analysis. *Archives of virology.* **145**: 1421-1436.

Bernal W., Smith H. M., Williams R. (1996). A community prevalence study of antibodies to Hepatitis A and E in inner-city London. *Journal of Medical Virology.* **49**: 230-234.

Bernuau J., Reuff B., Benhamou J. P. (1986). Fulminant and subfulminant liver failure: definitions and causes. *Seminars in Liver Disease.* **6**: 97-106.

Bi S. L., Purdy M. A., McCaustland K. A., Margolis H. S., Bradley D. W. (1993). The sequence of a hepatitis E virus isolated directly from a single source outbreak in China. *Virus Research.* **28**(3): 233-247.
(Genbank accession **M94177**)

Bissuel F., Houhou N., Leport C., Brun-Vezinet F., Vilde J-L. (1996). Hepatitis E antibodies and HIV status. *Lancet.* **347**: 1494.

Blumberg B. S., Gerstley B. S. J., Hungerford D. A., London W. T., and Stutnick A. J. (1967). A serum antigen (Australia antigen) in Down's syndrome, leukaemia, and hepatitis. *Annals of Internal Medicine.* **66**: 924-931.

Bradley D. W., Krawczynski K., Cook E. H., McCaustland K. A., Humphrey C. D., Spelbring J. E., Myint H., Maynard J. E. (1987). Enterically transmitted non-A, non-B hepatitis: Serial passage of disease in cynomolgus macaques and tamarins and recovery of disease-associated 27 to 34 nm viruslike particles. *Proceedings of the National Academy of Sciences USA.* **84**: 6277-6281.

Bradley D. W. (1992). Hepatitis E: Epidemiology, aetiology and molecular biology. *Reviews in Medical Virology.* **2**:19-28.

Bradley D. W. (1995). Hepatitis E virus: a brief review of the biology, molecular virology, and immunology of a novel virus. *Journal of Hepatology.* **22** (Suppl.1): 140-145.

Brillanti S., Garson J., Foli M., Whitby K., Deaville R., Masci C., Miglioloi M., Barbara L. (1994). A pilot study of combination therapy with ribavirin plus interferon alfa for alfa-resistant chronic hepatitis C. *Gastroenterology.* **107**: 812-817.

Bryan J. P., Tsarev S. A., Iqbal M., Ticehurst J., Emerson S., Ahmed A., Duncan J., Rafiqui A. R., Malik I. A., Purcell R. H., Legters L. J. (1994). Epidemic hepatitis E in Pakistan: patterns of serologic response and evidence that antibody to hepatitis E virus protects against disease. *Journal of Infectious Diseases.* **170**: 517-521.

Buisson Y., Coursaget P., Bercion R., Anne D., Debord T., Roue R. (1994). Hepatitis E virus infection in soldiers sent to endemic regions. *Lancet.* **344**: 1165-1166.

Buisson Y., Grandadam M., Nicand E., Cheval P., van Cuyck-Gandre H., Innis B., Rehel P., Coursaget P., Teyssou R., Tsarev S. (2000). Identification of a novel hepatitis E in Nigeria. *Journal of General Virology.* **81**:903-909.

Buti M., Jardi R., Cotrina M., Rodriguez-Frias F., Troonen H., Viladomiu L., Esteban J. I., Esteban R., Guardia J. (1995). Hepatitis E virus infection in acute hepatitis in Spain. *Journal of Virological Methods.* **55**: 49-54.

Buynak E. B., Roehm R. R., Tytell A. A., Bertland A. U., Lampson G. P., Hilleman M. R. (1976). Vaccine against human hepatitis B. *Journal of the American Medical Association.* **235**(26): 2832-2834.

Caggana M., Chan P., Ramsingh A. (1993). Identification of a single amino acid residue in the capsid of VP1 of coxsackievirus B4 that determines the virulent phenotype. *Journal of Virology* **67**: 4797-4803.

Caraceni P. and Van Thiel D. H. (1995). Acute liver failure. *Lancet.* **345**: 163-168.

Carman W. F., Jacyna M. R., Hadziyannis S., Karayiannis P., McGarvey M. A., Thomas H. C. (1989). Mutation preventing formation of e antigen in patients with chronic HBV infection. *Lancet.* **2**(8663): 588-591.

Chadha M. S., Walimbe A. M., Arankalle V. A. (1999). Retrospective serological analysis of hepatitis E patients: a long term follow-up study. *Journal of Viral Hepatitis* **6**(6): 457-61.

Charlton M., Adjei P., Poterucha J., Zein N., Moore B., Thernau T., Krom R., Wiesner R. (1998). TT-virus infection in North American blood donors, patients with fulminant hepatic failure, and cryptogenic cirrhosis. *Hepatology.* **28**: 839-842.

Chau K. H., Dawson, G. J., Bile K. M., Magnius L. O., Sjogren M. H., Mushawar I. K. (1993). Detection of IgA class antibody to hepatitis E virus in serum samples

from patients with hepatitis E virus infection. *Journal of Medical Virology.* **40**: 334-338.

Chauhan, A., Dilawari J. B., Jameel S., Kaur U., Chawla Y. K., Sharma M. L., Ganguly N. K. (1992). Common aetiological agent for epidemic and sporadic non-A, non-B hepatitis. *Lancet* **339**: 1509-1510.

Chauhan, A., Jameel S., Dilawari J. B., Chawla Y. K., Kaur U., Ganguly N. K. (1993). Hepatitis E virus transmission to a volunteer. *Lancet.* **341**: 149-150.

Chauhan A., Dilawari J. B., Kaur U., Ganguly N. K., Bushnurmath S., Chawla Y. K. (1994). Atypical strain of hepatitis E virus (HEV) from North India. *Journal of Medical Virology.* **44**: 22-29.

Chauhan A., Dilawari J. B., Sharma R., Mukesh M., Saroa S. R. (1998). Role of long-persisting human hepatitis E virus antibodies in protection. *Vaccine* **16**(7): 755-756.

Chobe L. P., Chadha M. S., Banerjee K., Arankalle V. A. (1997). Detection of HEV RNA in faeces, by RT-PCR during the epidemics of hepatitis E in India (1976-1995). *Journal of Viral Hepatitis.* **4**(2): 129-133.

Chomczynski P., Sacchi N. (1987). Single step method of RNA isolation by acid guanidinium thiocyanate-phenol-chloroform extraction. *Analytical Biochemistry.* **162**: 156-159.

Choo Q-L., Kuo G., Weiner A. J., Overby L. R., Bradley D. W., Houghton M. (1989). Isolation of a cDNA clone derived from a blood-borne non-A, non-B viral hepatitis genome. *Science.* **244**: 359-362.

Clayson E. T., Innis B. L., Myint K. S., Narupiti S., Vaughn D. W., Giri S., Ranabhat P., M. P. Shrestha. (1995a). Detection of hepatitis E virus infections among domestic swine in the Kathmandu Valley of Nepal. *American Journal of Tropical Medicine and Hygiene.* **53**(3): 228-232.

Clayson E. T., Myint K. S., Snitbhan R., Vaughn D. W., Innis B. L., Chan L., Cheung P., Shrestha M. P. (1995b). Viremia, fecal shedding, and IgM and IgG responses in patients with hepatitis E. *Journal of Infectious Diseases.* **172**(4): 927-933.

Clayson E. T., Shrestha M. P., Vaughn D. W., Snitbhan R., Shrestha K. B., Longer C. F., Innis B. L. (1997). Rates of hepatitis E virus infection and disease among adolescents and adults in Kathmandu, Nepal. *Journal of Infectious Diseases.* **176**(3): 763-766.

Cohen J. I., Ticehurst J. R., Purcell R. H., Buckler-White A., Baroudy B. M. (1987). Complete nucleotide sequence of wild-type hepatitis A virus: comparison with different strains of hepatitis A and other picornaviruses. *Journal of Virology.* **61**:55-59.

Corpet F. (1988). Multiple sequence alignment with hierarchical clustering. *Nucleic Acids Research.* **16**: 10881-10890.

Corwin A. L., Tien N. T. K., Bounlu K., Winarno J., Putri M. P., Laras K., Larasati R. P., Sukri N., Endy T., Sulaiman H. A., Hyams K. C. (1999). The unique riverine ecology of hepatitis E virus transmission in South-East Asia. *Transactions of the Royal Society of Tropical Medicine and Hygiene.* **93**(3): 255-260.

Coursaget P., Depril N., Sadi Yenen O., Cavuslu S., Badur S. (1993a). Hepatitis E virus infection in Turkey. *Lancet.* **342**: 810.

Coursaget P., Krawczynski K., Buisson Y., Nizou C., Molinie C. (1993b). Hepatitis E and hepatitis C virus infections among French soldiers with non-A, non-B hepatitis. *Journal of Medical Virology.* **39**:163-166.

Coursaget P., Depril N., Buisson Y., Molinie C., Roue R. (1994). Hepatitis type E in a French population - detection of anti-HEV by a synthetic peptide-based enzyme-linked immunosorbent assay. *Research in Virology.* **145**: 51-57.

Coursaget P., Buisson Y., N'Gawara M. N., Van Cuyck-Gandre H., Roue R. (1998). Role of hepatitis E virus in sporadic cases of acute and fulminant hepatitis in an endemic area (Chad). *American Journal of Tropical Medicine and Hygiene.* **58**(3): 330-334.

Courtney M. G., Omahoney M., Albloushi S., Sachithanandan S., Walshe J., Carmody M., Donoghue J., Parfrey N., Shattock A. G., Fielding J. (1994). Hepatitis E virus antibody prevalence. *Lancet.* 344:1166.

Dane D. S., Cameron C. H., Briggs M. (1970). Virus-like particles in serum of patients with Australia-antigen-associated hepatitis. *Lancet.* 1(7649): 695-698.

Darwish M. A., Faris R., Clemens J. D., Rao M. R., Edelman R. (1996). High seroprevalence of hepatitis A, B, C, and E viruses in residents in an Egyptian village in Nile Delta: A pilot study. *American Journal of Tropical Meicine and Hygiene.* 54(6):554-558.

Datta R., Panda S. K., Tandon B. N., Madnagopalan N., Bose S. L., Acharya S. K., Nayak N. C. (1987). Acute sporadic non-A, non-B viral hepatitis of adults in India, epidemiological and immunological studies. *Journal of Gastroenterology and Hepatology.* **2**: 333-345.

Davidson F., Simmonds P., Ferguson J. C., Jarvis L. M., Dow B. C., Follett E. A., Seed C. R., Krusius T., Lin C., Medgyesi G. A. (1995). Survey of major genotypes and subtypes of hepatitis C virus using RFLP of sequences amplified from the 5' non-coding region. *Journal of General Virology.* **76**: 1197-1204.

Dawson G. J., Mushahwar I. K., Chau K. H., and Gitnick G. L. (1992a). Detection of long-lasting antibody to hepatitis E virus in a US traveller to Pakistan [letter]. *Lancet* **340**: 426-427.

Dawson G. J., Chau K. H., Cabal C. M., Yarbough P. O., Reyes G. R., and Mushahwar I. K. (1992b). Solid-phase enzyme-linked immunosorbent assay for hepatitis E virus IgG and IgM antibodies utilizing recombinant antigens and synthetic peptides. *Journal of Virological Methods*. **38**: 175-186.

De Cock K. M., Bradley D. W., Sandford N. L., Govindarajan S., Maynard J. E., Redeker A. G. (1987). Epidemic non-A, non-B hepatitis in patients from Pakistan. *Annals of Internal Medicine*. **106**(2): 227-230.

Dear S., and Staden R. (1991). A sequence assembly program for efficient management of large projects. *Nucleic Acids Research*. **19**(14): 3907-3911.

Deinhardt F., Holmes A. W., Capps R. B., Popper H. (1967). Studies on the transmission of human viral hepatitis to marmoset monkeys. I. Transmission of disease, serial passages, and description of liver lesions. *Journal of Experimental Medicine*. **125**(4): 673-688.

Deka N., Sharma M. D., Mukerjee R. (1994). Isolation of the novel agent from human stool samples that is associated with sporadic non-A, non-B hepatitis. *Journal of Virology*. **6**: 7810-7815.

Domachowske J. B. and Rosenberg H. F. (1999). Respiratory syncytial virus infection: immune response, immunopathogenesis, and treatment. *Clinical Microbiology Reviews*. **12**(2): 298-309.

Donati M. C., Fagan E. A., Harrison T. J. (1997). Sequence analysis of full length HEV clones derived directly from human liver in fulminant hepatitis E. (in) Rizzetto M., Purcell R. H., Gerin J. L., Verme G. (Eds); *Viral Hepatitis and Liver Disease*: 313-316; Edizioni Minerva Medica, Torino (1997).
(Genbank accession **X98292**)

Doroshenko N. V., Smolina T. G., Zairov G. K., Stakhanova V. M. (1989). Serological proof of the existence of a variety of non-A, non-B hepatitis with a faecal-oral transmission mechanism. *Voprosy Virusologii*. **34**(2): 164-167.

Durand F., Degott C., Sauvanet A., Molas G., Sicot C., Marcellin P., Belghiti J., Erlinger S., Benhamou J. P., Bernuau J. (1997). Subfulminant syncytial giant cell hepatitis: recurrence after liver transplantation treated with ribavirin. *Journal of Hepatology*. **26**: 722-726.

Dusheiko G. M., Khakoo S., Soni P., Grellier L. (1996). A rational approach to the management of hepatitis C infection. *British Medical Journal*. **312**: 357-364.

el-Zimaity D. M., Hyams K. C., Imam I. Z., Watts D. M., Bassily S., Naffea E. K., Emara K., Burans J., Purdy M. A. (1993). Acute sporadic hepatitis E in an Egyptian pediatric population. *American Journal of Tropical Medicine & Hygiene*. **48**: 372-376.

Erker J. C., Desai S. M., Schlauder G. G., Dawson G. J., Mushahwar I. K. (1999). A hepatitis E virus variant from the United States: molecular characterization and transmission in cynomolgus macaques. *Journal of General Virology*. **80**: 681-690.

Fabrizi F., Lunghi G., Bacchini G., Corti M., Pagano A., Locatelli F. (1997). Hepatitis E virus infection in haemodialysis patients: a seroepidemiological survey. *Nephrol. Dial. Transplant*. **12**(1): 133-136.

Fagan E. A., Ellis D. S., Tovey G. M., Lloyd G., Smith H. M., Portmann B., Tan K. C., Zuckerman A. J., Williams R. (1992). Togavirus-like particles in acute liver failure attributed to sporadic non-A, non-B hepatitis and recurrence after liver transplantation. *Journal of Medical Virology*. **38**: 71-77.

Fagan E. A., Menon T., Valliammai T., Paul D. A., Donati M. C., Burroughs A. K., Gunamanam D., Dhevahi E., Harrison T. J. (1994). Equivocal serological diagnosis of sporadic fulminant hepatitis E in pregnant Indians. *Lancet*. **344**: 342-343.

Favorov M. O., Fields H. A., Purdy M. A., Yashina T. L., Aleksandrov A. G., Alter M. J., Yarasheva D. M., Bradley D. W., Margolis H. S. (1992). Serologic identification of hepatitis E virus infections in epidemic and endemic settings. *Journal of Medical Virology*. **36**: 246-250.

Favorov M. O., Khudyakov Y. E., Fields H. A., Khudyakova N. S., Padhye N., Alter M. J., Mast E., Polish L., Yashina T. L., Yarasheva D. M., Onischenko G. G., Margolis H. S. (1994). Enzyme immunoassay for the detection of antibody to hepatitis E virus based on synthetic peptides. *Journal of Virological Methods*. **46**: 237-250.

Favorov M. O., Khudyakov Y. E., Mast E. E., Yashina T. L., Shapiro C. N., Khudyakova N. S., Jue D. L., Onischenko G. G., Margolis H. S., Fields H. A. (1996). IgM and IgG antibodies to hepatitis E virus (HEV) detected by an enzyme immunoassay based on an HEV-specific artificial recombinant mosaic protein. *Journal of Medical Virology*. **50**(1): 50-58.

Favorov M. O., Kosoy M. Y., Tsarev S. A., Childs J. E., Margolis H. S. (2000). Prevalence of antibody to hepatitis E virus among rodents in the United States. *Journal of Infectious Diseases*. **181**: 449-55.

Feinberg A. P., and Vogelstein B. (1983). A technique for radiolabelling DNA restriction endonuclease fragments to high specific activity. *Anals of Biochemistry*. **132**: 6-13.

Fimmel C. J., Guo L., Compans R. W., Brunt E. M., Hickman S., Perrillo R. R., Mason A. L. (1998). A case of syncytial giant cell hepatitis with features of paramyxoviral infection. *American Journal of Gastroenterology*. **93**: 1931-1937.

Focaccia R., Sette H., and Conceicao O. J. G. (1995). Hepatitis E in brazil. *Lancet*. **346**: 1165.

Frank C., Mohamed M. K., Strickland G. T., Lavanchy D., Arthur R. R., Magder L. S., El Khoby T., Abdel-Wahab Y., Ohn S. A., Anwar W., Sallam I. (2000). The role of parenteral antischistosomal therapy in the spread of hepatitis C virus in Egypt. *Lancet.* **355**: 887-891.

Galibert F., Mandart E., Fitoussi F., Tiollais P., Charnay P. (1979). Nucleotide sequence of the hepatitis B virus genome (subtype ayw) cloned in *E. coli. Nature.* **281**: 646-650.

Gerin J. L. (1994). The taxonomy of hepatitis delta virus. *Viral Hepatitis and Liver Disease.* 1994: pp63-64.

Garkavenko O., Anderson D., Schroeder B., Croxson C. (2000). Evidence of swine hepatitis virus in New Zealand. *Antiviral Therapy.* **5**: E10.

Gauntt, C. J., and Pallansch M. A. (1996). Coxsackievirus B3 clinical isolates and murine myocarditis. *Virus Research* **41**: 89-99.

Ghabrah T. M., Tsarev S., Yarbough P. O., Emerson S. U., Strickland G. T., Purcell R. H. (1998). Comparison of tests for antibody to hepatitis E virus. *Journal of Medical Virology.* **55**(2): 134-137.

Gimson A. E., O'Grady J., Ede R. J., Portmann B., Williams R. (1986). Late-onset hepatic failure: clinical, serological and histological features. *Hepatology* **6**: 188-194.

Goldsmith R., Yarbough P. O., Reyes G. R., Fry K. E., Gabor K. A., Kamel M., Zakaria S., Amer S., Gaffar Y. (1992). Enzyme-linked immunosorbent assay for diagnosis of acute sporadic hepatitis E in Egyptian children. *Lancet.* **339**: 328-331.

Gouvea V., Hoke C. H., Innis B. L. (1998). Genotyping of hepatitis E virus in clinical specimens by restriction endonuclease analysis. *Journal of Virological Methods.* **70**(1): 71-78.

Green K. Y., Ando T., Balayan M. S., Berke T., Clarke I. N., Estes M. K., Thiel H. J. (2000). Taxonomy of the caliciviruses. *Journal of Infectious Diseases.* 181(suppl2):S322-330.

Gupta D. N., and Smetana H. F. (1957). The histopathology of viral hepatitis as seen in Delhi epidemic (1955-56). *Indian Journal of Medical Research* (suppl). **45**: 101-113.

Gust I. D., and Purcell R. H. (1987). Report of a workshop: Waterborne non-A, non-B hepatitis. *Journal of Infectious Diseases.* **156**(4): 630-635.

Hadziyannis S. J. (1997). Review: hepatitis delta. *Journal of Gastroenterology and Hepatology.* **12**(4): 289-298.

Halbur P. G., Kasorndorkbua C., Gilbert C., Guenette D., Potters M. B., Purcell R. H., Emerson S. U., Toth T. E., Meng X. J. (2001). Comparative pathogenesis of

infection of pigs with hepatitis E viruses recovered from a pig and a human. *Journal of Clinical Microbiology*. **39**(3): 918-923.

Halfon P., Ouzan D., Chanas M., Khiri H., Feryn J. M., Mangin L., Masseyef M. F., Salvadori J. M. (1994). High prevalence of hepatitis E virus antibody in haemodialysis patients. *Lancet*. **344**: 746.

Halstead S. B. (1988). Pathogenesis of Dengue: challenges to molecular biology. *Science*. **239**: 476-481.

Harrison T. J., Hopes E. A., Oon C. J.,Zanetti A. R., Zuckerman A. J. (1991). Independent emergence of a vaccine-induced escape mutant of hepatitis B virus. *Journal of Hepatology*. **13**: S105-S107.

He J. K., Ching W. M., Yarbough P., Wang H., Carl M. (1995). Purification of a baculovirus-expressed hepatitis E virus structural protein and utility in an enzyme-linked immunosorbent assay. *Journal of Clinical Microbiology*. **33**(12): 3308-3311.

He J. K., Hoffman S. L., and Hayes C. G. (1997). DNA inoculation with a plasmid vector carrying the hepatitis E virus structural protein gene induces immune response in mice. *Vaccine*. **15**(4): 357-362.

He J., Hayes C. G., Binn L. N., Seriwatana J., Vaughn D. W., Kuschner R. A., Innis B. L. (2001). Hepatitis E virus DNA vaccine elicits immunologic memory in mice. *Journal of Biomedical Science*. **8**(2): 223-226.

Hijikata M., Takahashi K., and Mishiro S. (1999). Complete circular DNA genome of a TT virus variant (isolate name SANBAN) and 44 partial ORF2 sequences implicating a great degree of diversity beyond genotypes. *Virology*. **260**: 17-22.

Hsieh S-Y., Meng X-J., Wu Y-H., Liu S-T., Tam A. W., Lin D-Y., Liaw Y-F. (1999). Identity of a novel swine hepatitis E virus in Taiwan forming a monophyletic group with Taiwan isolates of human hepatitis E. *Journal of Clinical Microbiology*. **37**: 3828-3834.

Huang C. C., Nguyen D., Fernandez J., Yun K. Y., Fry K. E., Bradley D. W., Tam A. W., Reyes G. R. (1992). Molecular cloning and sequencing of the Mexico isolate of hepatitis E virus (HEV). *Virology*. **191**: 550-558.
(Genbank accession **M74506**)

Huang R. T., Li D. R., Wei J., Huang X. R., Yuan X. T., Tian X. (1992). Isolation and identification of hepatitis E virus in Xinjiang, China. *Journal of General Virology*. **73**: 1143-1148.

Huang R. T., Nakazono N., Ishii K., Li D. R., Kawamata O., Kawaguchi R., Tsukada Y. (1995). Hepatitis E virus (87A strain) propagated in A549 cells. *Journal of Medical Virology*. **47**(4): 299-302.

Huang R. T., Paxton W. A., Wolinsky S. M., Neumann A. U., Zhang L., He T., Kang S., Ceradini D., Jin Z., Yazdanbakhsh K., Kunstman K., Erickson D.,

Dragon E., Landau N. R., Phair J., Ho D. D., Koup R. A. (1997). The role of mutant CCR5 allele in HIV-1 transmission and disease progression. *Nature Medicine.* 3(3): 252-253.

Hyams K. C., McCarthy M. C., Kaur M., Purdy M. A., Bradley D. W., Mansour M. M., Gray S., Watts D. M., Carl M. (1992). Acute sporadic hepatitis E in children living in Cairo, Egypt. *Journal of Medical Virology.* 37: 274-277.

Hyams K. C., Purdy M. A., Kaur M., McCarthy M. C., Hussain M. A., El-Tigani A., Krawczynski K., Bradley D. W., Carl M. (1992). Acute sporadic hepatitis E in Sudanese children: analysis based on a new western blot assay. *Journal of Infectious Diseases.* 165: 1001-1005.

Ibarra H. V., Riedemann S. G., Siegel F. G., Reinhardt G. V., Toledo C. A., Frosner G. (1994). Hepatitis E virus in chile. *Lancet.* 344: 1501.

Ikeda H., Takasu M., Inoue K., Okamato H., Miyakawa Y., Mayumi M. (1999). Infection with an unenveloped DNA virus (TTV) in patients with acute or chronic liver disease of unknown of unknown etiology and in those positive for hepatitis C virus RNA. *Journal of Hepatology.* 30: 205-212.

Inoue K., Yoshiba M., Sekiyama K., Okamoto H., Mayumi M. (1998). Clinical and molecular virological differences between fulminant hepatic failures following acute and chronic infection with hepatitis B virus. *Journal of Medical Virology.* 55(1): 35-41.

Inoue O., Nagataki S., Itakura H., Shimizu Y., Yano M., Shikata T., Tandon B.N. (1986). Liver morphology in marmosets infected with epidemic non-A, non-B hepatitis in India. *Liver.* 6: 178-183.

Iqbal M., Ahmed A., Qamar A., Dixon K., Duncan J.F., Islam N.U., Rauf A., Bryan J.P., Malik I.A., Legters L.J. (1989). An outbreak of enterically transmitted non-A, non-B hepatitis in Pakistan. *American Journal of Tropical Medicine and Hygiene.* 40: 438-443.

Jaiswal S. P., Jain A. K., Naik G., Soni N., Chitnis D. S. (2001). Viral hepatitis in pregnancy. *International Journal of Gynaecology and Obstetrics.* 72(2): 103-108.

Jameel S., Durgapal H., Habibullah C. M., Khuroo M. S., Panda S. K. (1993). Enteric non-A, non-B hepatitis: epidemics, animal transmission and hepatitis E virus detection by the polymerase chain reaction. *Journal of Medical Virology.* 37: 263-270.

Jameel S., Zafrullah M., Ozdener M. H., Panda S. K. (1996). Expression in animal cells and characterization of the hepatitis E virus structural proteins. *Journal of Virology.* 70(1): 207-216.

Jardi R., Buti M., Rodriguez-Frias F., R. Esteban. (1993). Hepatitis E infection in acute sporadic hepatitis in Spain. *Lancet.* 341:1355-1356.

Jiang X., Wang M., Graham D. Y., Estes M. K. (1992). Expression, self-assembly and antigenicity of the Norwalk virus capsid protein. *Journal of Virology.* **66**: 6527-6532.

Joshi Y. K., Babu S., Sarin S., Tandon B. N., Gandhi B. M., Chaturvedi V. C. (1985). Immunoprophylaxis of epidemic non-A, non-B hepatitis. *Indian Journal of Medical Research.* **81**: 18-19.

Jukes T. H., and Cantor C. R. (1969). Evolution of protein molecules. In: Munro H. H. (ed.), *Mammalian protein metabolism.* Academic Press, New York, pp. 21-132.

Kabrane-Lazizi Y., Meng X-J., Purcell R. H., Emerson S. U. (1999a). Evidence that the genomic RNA of hepatitis E virus is capped. *Journal of Virology.* **73**(10): 8848-8850.

Kabrane-Lazizi Y., Fine J. B., Elm J., Glass G. E., Higa H., Diwan A., Gibbs C. J. Jr., Meng X. J., Emerson S. U., Purcell R. H. (1999b). Evidence for widespread infection of wild rats with hepatitis E virus in the United States. *American Journal of Tropical Medicine and Hygiene.* **61**(2): 331-335.

Kabrane-Lazizi Y., Zhang M., Purcell R. H., Miller K. D., Davey R. T., Emerson S. U. (2001). Acute hepatitis caused by a novel strain of hepatitis E virus most closely related to United States strains. *Journal of General Virology.* **82**: 1687-1693.

Kamel M. A., Troonen H., Kapprell H. P., ElAyady A., Miller F. D. (1995). Seroepidemiology of hepatitis E virus in the Egyptian Nile Delta. *Journal of Medical Virology.* **47**(4): 399-403.

Kane M. A., Bradley D. W., Shrestha S. M., Maynard J. E., Cook E. H., Mishra R. P., Joshi D. D. (1984). Epidemic non-A, non-B hepatitis in Nepal. Recovery of a possible etiologic agent and transmission studies in marmosets. *Journal of the American Medical Association.* **252**(22): 3140-3145.

Karetnyi Y. V., Favorov M. O., Khudyakova N. S., Weiss P., Bar-shani S., Handsher R., Aboudy Y., Varsano N., Schwartz E., Levin E., Mendelson E., Fields H. A. (1995). Serological evidence for hepatitis E virus infection in Israel. *Journal of Medical Virology.* **45**: 316-320.

Karetnyi Y. V., Gilchrist M. J., Naides S. J. (1999). Hepatitis E virus infection prevalence among selected populations in Iowa. *Journal of Clinical Virology.* **14**(1): 51-55.

Kasakura S. (1971). A factor in maternal plasma during pregnancy that suppresses the reactivity of mixed lymphocyte cultures. *Journal of Immunology.* **107**: 1296-1301.

Kao J-H., Chen W., Chen P-J., Lai M-Y., Chen D-S. (2000). TT virus infection in patients with chronic hepatitis B or C: Influence on clinical, histological and virological features. *Journal of Medical Virology.* **60**: 387-392.

Kaufman B., Gandhi S. A., Louie E., Rizzi R., Illei P. (1997). Herpes simplex virus hepatitis: case report and review. *Clinical Infectious Diseases.* **24**(3): 334-8.

Kaur M., Hyams K. C., Purdy M. A., Krawczynski K., Ching W. M., Fry K. E., Reyes G. R., Bradley D. W., Carl M. (1992). Human linear B-cell epitopes encoded by the hepatitis E virus include determinants in the RNA-dependent RNA polymerase. *Proceedings of the National Academy of Science USA.* **89**: 3855-3858.

Kawai H. F., Koji T., Iida F., Kaneko S., Kobayashi K., Nakane P. K. (1999). Shift of hepatitis E RNA from hepatocytes to biliary epithelial cells during acute infection of rhesus monkey. *Journal of Viral Hepatitis.* **6**: 287-297.

Kazachkov Y. A., Balayan M. S., Ivannikova T. A., Panina L. I., Orlova T. M., Zamyatina N. A., Kusov Y. Y. (1992). Hepatitis E virus in cultivated cells. *J.Arch. Virol.* **127**: 399-402.

Ke W. M., Tan D., Li J.G., Izumi S., Shinji Y., Hotta H., Yao J. L. (1996). Consecutive evaluation of immunoglobulin M and G antibodies against hepatitis E virus. *Journal of Gastroenterology.* **31**(6): 818-822.

Khudyakov Y. E., Favorov M. O., Jue D. L., Hine T. K., Fields H. A. (1994a). Immunodominant antigenic regions in a structural protein of the hepatitis E virus. *Virology.* **198**: 390-393.

Khudyakov Y. E., Favorov M. O., Khudyakova N. S., Cong M., Holloway B. P., Padhye N., Lambert S. B., Jue D. L., Fields H. A. (1994b). Artificial mosaic protein containing antigenic epitopes of hepatitis E virus. *Journal of Virology.* **68**: 7067-7074.

Khudyakov Yu. E., Lopareva E. N., Jue D. L., Crews T. K., Thyagarajan S. P., Fields H. A. (1999). Antigenic domains of the open reading frame 2-encoded protein of hepatitis E virus. *Journal of Clinical Microbiology.* **37**(9): 2863-2871.

Khuroo M. S. (1980). Study of an epidemic of non-A, non-B hepatitis. Possibility of another human hepatitis virus distinct from post-transfusion non-A, non-B type. *American Journal of Medicine.* **68**(6): 818-824.

Khuroo M. S., Teli M. R., Skidmore S., Sofi M. A., Khuroo M. I. (1981). Incidence and severity of viral hepatitis in pregnancy. *American Journal of Medicine.* **70**(2): 252-255.

Khuroo M. S., Duermeyer W., Zargar S. A., Ahanger M. A., Shah M. A. (1983). Acute sporadic non-A, non-B hepatitis in India. *American Journal of Epidemiology.* **118**(3): 360-364.

Khuroo M. S , Dar M. Y. (1992). Hepatitis E: evidence for person to person transmission and inability of low dose immune serum globulin from an Indian source to prevent it. *Indian Journal of Gastroenterology.* **11**(3): 113-116.

Khuroo M. S., Kamili S., Dar M. Y., Moecklii R., Jameel S. (1993). Hepatitis E and long-term antibody status. *Lancet.* **341**:1355.

Khuroo M. S., Rustgi V. K., Dawson G. J., Mushahwar I. K., Yattoo G. N., Kamili S., Khan B. A.(1994). Spectrum of hepatitis E virus infection in India. *Journal of Medical Virology.* **43**: 281-286.

Khuroo M. S., Kamili S., Jameel S. (1995). Vertical transmission of hepatitis E virus. *Lancet.* **345**: 1025-1026.

Koonin E. V., Gorbalenya A. E., Purdy M. A., Rozanov M. N., Reyes G. R., Bradley D. W. (1992). Computer-assisted assignment of functional domains in the nonstructural polyprotein of hepatitis E virus: delineation of an additional group of positive-strand RNA plant and animal viruses. *Proceedings of the National Academy of Science USA.* **89**: 8259-8263.

Koshy A., Richards A. L., Al-Mufti S., Grover S., Shabrawy M. A., Pacsa A., Al-Anezi A. A. H., Al-Nakib B., Burans J., Carl M., Hyams K. C. (1994). Acute sporadic hepatitis-E in Kuwait. *Journal of Medical Virology.* **42**: 405-408.

Krawczynski K., and Bradley D. W. (1989). Enterically transmitted non-A, non-B hepatitis: Identification of virus-associated antigen in experimentally infected cynomolgus macaques. *Journal of Infectious Diseases.* **159**: 1042-1049.

Krawczynski K. (1993). Hepatitis E. *Hepatology.* **17**: 932-941.

Kwo P. Y., Schlauder G. G., Carpenter H. A., Murphy P. J., Rosenblatt J. E., Dawson G. J., Mast E. E., Krawczynski K., Balan V. (1997). Acute hepatitis E by a new isolate acquired in the United States. *Mayo Clinic Proceedings.* **72**(12): 1133-1136.

Lau D. T., Everhart J., Kleiner D. E., Park Y., Vergalla J., Schmid P., Hoofnagle J. H. (1997). Long-term follow-up of patients with chronic hepatitis B treated with interferon alfa. *Gastroenterology.* **113**(5): 1660-1667.

Lau D. T., Farooq Khokhar M., Doo E., Ghany M. G., Herion D., Park Y., Kleiner D. E., Hoofnagle J. H. (2000). Long-term therapy of chronic hepatitis B with lamivudine. *Hepatology.* **32**: 828-834.

Lau J. Y. N., Sallie R., Fang J. W. S., Yarbough P. O., Reyes G. R., Portmann B. C., Mieli-Vergani G., Williams R. (1995). Detection of hepatitis E virus genome and gene products in two patients with fulminant hepatitis E. *Journal of Hepatology.* **22**: 605-610.

Laurent S., Vautherot J-F., Madelaine M-F., Gall G. L., Rasschaert D. (1994). Recombinant rabbit haemorrhagic disease virus capsid protein expressed in Baculovirus self-assembles into virus-like particles and induces protection. *Journal of Virology.* **68**(10): 6794-6798.

Lavanchy D., Morel B., Frei P. C. (1994). Seroprevalence of hepatitis E virus in Switzerland. *Lancet.* **344**: 747-748.

Lee S. D., Wang Y. J., Lu R. H., Chan C. Y., Lo K. J., Moeckli R. (1994). Seroprevalence of antibody to hepatitis E virus among chinese subjects in Taiwan. *Hepatology.* **19**: 866-870.

Li F., Zhuang H., Kolivas S., Locarnini S. A., Anderson D. A. (1994). Persistent and transient antibody responses to hepatitis E virus detected by western immunoblot using open reading frame 2 and 3 and glutathione s-transferase fusion proteins. *Journal of Clinical Microbiology.* **32**: 2060-2066.

Li F., Riddell M. A., Seow H-F., Takeda N., Miyamura T., Anderson D. A. (2000). Recombinant subunit ORF2.1 antigen and induction of antibody against immunodominant epitopes in the hepatitis E virus capsid protein. *Journal of Medical Virology.* **60**: 379-386.

Li T. C., Yamakawa Y., Suzuki K., Tatsumi M., Razak M. A., Uchida T., Takeda N., Miyamura T. (1997). Expression and self-assembly of empty virus-like particles of hepatitis E virus. *Journal of Virology.* **71**(10): 7207-7213.

Lindh M., Gonzalez J. E., Norkrans G., Horal P. (1998). Genotyping of hepatitis B virus by restriction pattern analysis of a pre-S amplicon. *Journal of Virological Methods.* **72**(2):163-174.

Lindsay K. L., Trepo C., Heintges T., Shiffman M. L., Gordon S. C., Hoefs J. C., Goodman Z. D., Laughlin M., Yao R., Albrecht J. K.; The Hepatitis Intervention Therapy Group. (2001). A randomized, double-blind trial comparing pegylated interferon alfa-2b to interferon alfa-2b as initial treatment for chronic hepatitis C. *Hepatology.* **34**(2): 395-403.

Ling R., and Harrison T. J. (1999). Functional analysis of mutations conferring lamivudine resistance on hepatitis B virus. *Journal of General Virology.* **80**(3): 601-606.

Linnen J., Wages Jr J., Zhang-Keck Z-Z., Fry K. E., Krawczynski K. Z., Alter H., Koonin E., Gallagher M., Alter M., Hadziyannis S., Karayiannis P., Fung K., Nakatsuji Y., Shih J. W., Young L., Piatak M. Jr, Hoover C., Fernandez J., Chen S., Zou J. C., Morris T., Hyams K. C., Ismay S., Lifson J. D., Kim J. P., (1996). Molecular cloning and disease association of Hepatitis G virus: A transfusion-transmissible agent. *Science.* **271**: 505-508.

Lok A. S., Kwan W. K., Moeckli R., Yarbough P. O., Chan R. T., Reyes G. R., Lai C. L., Chung H-T., Lai T. S. T. (1992). Seroepidemiological survey of hepatitis E in Hong Kong by recombinant-based enzyme immunoassays. *Lancet.* **340**: 1205-1208.

Lok A. S. F. and C. Soldevila-Pico. (1994). Epidemiology and serologic diagnosis of hepatitis E. *Journal of Hepatology.* **20**: 567-569.

Longer C. F., Denny S. L., Caudill J. D., Miele T. A., Asher L. V., Myint K. S., Huang C-C., Engler W. F., LeDuc J. W., Binn L. N., Ticehurst J. R. (1993).

Experimental hepatitis E: pathogenesis in cynomolgus macaques (*Macaca fascicularis*). *Journal of Infectious Diseases.* **168**: 602-609.

McAleer W. J., Buynak E. B., Maigetter R. Z., Wampler D. E., Miller W. J., Hilleman M. R. (1984). Human hepatitis B vaccine from recombinant yeast. *Nature.* **307**: 178-180.

McCrudden R., O'Connell S., Farrant T., Beaton S., Iredale J. P., Fine D. (2000). Sporadic acute hepatitis E in the United Kingdom: an underdiagnosed phenomenon? *Gut.* **46**: 732-733.

Magden J., Takeda N., Li T., Auvinen P., Ahola T., Miyamura T., Merits A., Kaariainen L. (2001). Virus-specific mRNA capping enzyme encoded by hepatitis E virus. *Journal of Virology.* **75**(14): 6249-6255.

Malathi S., Mohanavalli B., Menon T., Srilatha P., Sankaranarayanan V. S., Raju B. B., Ramathilagan B., Thyagarajan S. P. (1998). Clinical and viral markers pattern of acute sporadic hepatitis in children in Madras, South India. *Journal of Tropical Pediatrics.* **44**(5): 275-278.

Maneerat Y., Clayson E. T., Myint K. S., Young G. D., Innis B. L. (1996). Experimental infection of the laboratory rat with the hepatitis E virus. *Journal of Medical Virology.* **48**(2): 121-128.

Mannucci P. M., Gringeri A., Sàntagostino E., Romano L., Zanetti A. (1994). Low risk of transmission of hepatitis E virus by large-pool coagulation factor concentrates. *Lancet.* **343**: 597-598.

Mast E. E., Kuramoto I. K., Favorov M. O., Schoening V. R., Burkholder B. T., Shapiro C. N., Holland P. V. (1997). Prevalence of and risk factors for antibody to hepatitis E virus seroreactivity among blood donors in northern California. *Journal of Infectious Diseases.* **176**(1): 34-40.

Mast E. E., Alter M. J., and Holland P. V. (1998). Evaluation of assays for antibody to hepatitis E virus by a serum panel. *Hepatology.* **27**(3): 857-861.

Mateos M. L., Camarero C., Lasa E., Teruel J. L., Mir N., Baquero F. (1999). Hepatitis E virus: relevance in blood donors and risk groups. *Vox Sanguinis.* **76**(2): 78-80.

Medhat A., el-Sharkawy M. M., Shaaban M.M., Ghaneima S. E. (1993). Acute viral hepatitis in pregnancy. *International Journal of Gynaecology and Obstetrics.* **40**(1): 25-31.

Medrano F. J., Sanchez-Quijano A., Torronteras R., Leal M., Lissen E. (1994). Hepatitis E virus and HIV infection in homosexual men. *Lancet.* **345**: 127.

Melnick J. L. (1992). Properties and classification of hepatitis A virus. *Vaccine.* **10** (suppl 1): S24-26.

Meng X. J., Purcell R. H., Halbur P. G., Lehman J. R., Webb D. M., Tsareva T. S., Haynes J. S., Thacker B. J., Emerson S. U. (1997). A novel virus in swine is closely related to the human hepatitis E virus. *Proceedings of the National Academy of Sciences USA.* **94**(18): 9860-9865.

Meng X. J., Halbur P. G., Haynes J. S., Tsareva T. S., Bruna J. D., Royer R. L., Purcell R. H., Emerson S. U. (1998a). Experimental infection of pigs with the newly identified swine hepatitis E virus (swine HEV), but not with human strains of HEV. *Archives of Virology.* **143**(7): 1405-1415.

Meng J., Pillot J., Dai X., Fields H. A., Khudyakov Y. E. (1998b). Neutralization of different geographic strains of the hepatitis E virus with anti-hepatitis E virus-positive serum samples obtained from different sources. *Virology.* **249**(2): 316-324.

Meng X. J., Halbur P. G., Shapiro M. S., Govindarajan S., Bruna J. D., Mushahwar I. K., Emerson S. U. (1998c). Genetic and experimental evidence for cross-species infection by swine hepatitis E virus. *Journal of Virology.* **72**(12): 9714-9721.

Meng X. J., Dea S., Engle R. E., Friendship R., Lyoo Y. S., Sirinarumitr T., Emerson S. U. (1999). Prevalence of antibodies to the hepatitis E virus in pigs from countries where hepatitis E is common or is rare in the human population. *Journal of Medical Virology.* **59**: 297-302.

Minor P.D. (1992). The molecular biology of poliovaccines. *Journal of General Virology.* **73**: 3065-3077.

Moaven L., Van Asten M., Crofts N., Locarnini S. A. (1995). Seroepidemiology of hepatitis E in selected Australian populations. *Journal of Medical Virology.* **45**: 326-330.

Montella F., Rezza G., Di Sora F., Pezzotti P., Recchia O. (1994). Association between hepatitis E virus and HIV infection in homosexual men. *Lancet.* **344**: 1433.

Mushahwar I. K., Dawson G. J., Bile K. M., Magnius L. O. (1993). Serological studies of an enterically transmitted non-A, non-B hepatitis in Somalia. *Journal of Medical Virology.* **40**: 218-221.

Mushahwar I. K., Erker J. C., Muerhoff A. S., Leary T. P., Simons J. N., Birkenmeyer L. G., Chalmers M. L., Pilot-Matias T. J., Dexai S. M. (1999). Molecular and biophysical characterization of TT virus: evidence for a new virus family infecting humans. *Proceedings of the National Academy of Sciences USA.* **96**: 3177-3182.

Mushahwar I. K. (2000). Recently discovered blood-borne viruses: are they hepatitis viruses or merely endosymbionts? *Journal of Medical Virology.* **62**: 399-404.

Myint H., Soe M. M., Khin T., Myint T. M., Tin K. M. (1985). A clinical and epidemiological study of an epidemic of non-A, non-B hepatitis in Rangoon. *American Journal of Tropical Medicine and Hygiene.* **34**: 1183-1189.

Naik S. R., Aggarwal R., Salunke P. N., Mehrotra N. N. (1992). A large waterborne viral hepatitis E epidemic in Kanpur, India. *Bulletin World Health Organisation.* **70**: 597-604.

Nanda S. K., Panda S. K., Durgapal H., Jameel S. (1994). Detection of the negative strand of hepatitis E virus RNA in the livers of experimentally infected rhesus monkeys - evidence for viral replication. *Journal of Medical Virology.* **42**: 237-240.

Nanda S. K., Yalcinkaya K., Panigrahi A. K., Acharya S. K., Jameel S., Panda. S. K. (1994). Etiological role of hepatitis E virus in sporadic fulminant hepatitis. *Journal of Medical Virology.* **42**: 133-137.

Nanda S. K., Ansari I. H., Acharya S. K., Jameel S., Panda S. K. (1995). Protracted viremia during acute sporadic hepatitis E virus infection. *Gastroenterology.* **108**: 225-230.

Nandwani R. (1995). Hepatitis E virus and HIV infection in homosexual men. *Lancet.* **345**: 126-127.

Naoumov N. V., Petrova E. P., Thomas M. G., Williams R. (1998). Presence of a newly described human DNA virus (TTV) in patients with liver disease. *Lancet.* **352**: 195-197.

Nayak N. C., Panda S. K., Datta R., Zuckerman A. J., Guha D. K., Madanagopalan N., Buckshee K. (1989). Aetiology and outcome of acute viral hepatitis in pregnancy. *Journal of Gastroenterology and Hepatology.* **4**: 345-352.

Nishizawa T., Okamoto H., Konishi K., Yoshizawa H., Miyakawa Y., Mayumi M. (1997). A novel DNA virus (TTV) associated with elevated transaminase levels in posttransfusion hepatitis of unknown etiology. *Biochem Biophys Res Commun.* **241**(1): 92-97.

Norder H., Courouce A-M., Magnius L.O. (1994). Complete genomes, phylogenetic relatedness, and structural proteins of six strains of the Hepatitis B virus, four of which represent two new genotypes. *Virology.* **198**: 489-503.

Okamoto H., Akahane Y., Ukita M., Fukuda M., Tsuda F., Miyakawa Y., Mayumi M. (1998). Fecal excretion of a nonenveloped DNA virus (TTV) associated with posttransfusion non-A-G hepatitis. *Journal of Medical Virology.* **56**: 128-132.

Panda S. K., Datta R., Kaur J., Zuckerman A. J., Nayak N. C. (1989). Enterically transmitted non-A, non-B hepatitis: recovery of virus-like particles from an epidemic in south Delhi and transmission studies in rhesus monkey. *Hepatology.* **10**: 466-472.

Panda S. K., Nanda S. K., Zafrullah M., Ansari I. U. H., Ozdener M. H., Jameel S. (1995). An Indian strain of hepatitis E virus (HEV): cloning, sequence, and

expression of structural region and antibody responses in sera from individuals from an area of high-level HEV endemicity. *Journal of Clinical Microbiology*. **33**: 2653-2659.

Panda S. K., Ansari I. H., Durgapal H., Jameel S. (2000). The *in-vitro* synthesized RNA from a cDNA clone of hepatitis E virus is infectious. *Journal of Virology*. **74**(5): 2430-2437.
(Genbank accession **AF076239**)

Paul D. A., Knigge M. F., Ritter A., Gutierrez R., Pilot-Matias T., Chau K. H., Dawson G. J. (1994). Determination of hepatitis E virus seroprevalence by using recombinant fusion proteins and synthetic peptides. *Journal of Infectious Diseases*. **169**: 801-806.

Pawlotsky J. M., Belec L., Gresenguet G., Deforges L., Bouvier M., Duval J., Dhumeaux D. (1995). High prevalence of hepatitis B, C, and E markers in young sexually active adults from the Central African Republic. *Journal of Medical Virology*. **46**: 269-273.

Pham B. N., Guigonis V., Durand F., Bernuau J., Udin L., Bezeaud A. (1994). Anti-HEV antibodies in acute hepatitis in France (letter). *Journal of Hepatology*. **20**: 679.

Phillips M. J., Blendis L. M., Poucell S., Offterson J., Petric M., Roberts E., Levy G. A., Superina R. A., Greig P. D., Cameron R. (1991). Syncytial giant-cell hepatitis. Sporadic hepatitis with distinctive pathological features, a severe clinical course, and paramyxoviral features. *New England Journal of Medicine*. **324**: 455-460.

Pillot J., Meng J., Maillard P., Dauguet C. (1997). The ''presumed'' hepatitis F virus. In: Viral Hepatitis and Liver Disease. Editors Rizzetto M., Purcell R. H., Gerin J. L. and Verme G. Edizioni Minerva Medica, Turin. pp399-401.

Pina S., Jofre J., Emerson S. U., Purcell R. H., Girones R. (1998). Characterization of a strain of infectious hepatitis E virus isolated from sewage in an area where hepatitis E is not endemic. *Appl Environ Microbiol*. **64**(11):4485-4488.

Pina S., Buti M., Cotrina M., Piella J., Girones R. (2000). HEV identified in serum from humans with acute hepatitis and in sewage of animal origin in Spain. *Journal of Hepatology*. **33**: 826-833.

Pisanti F. A., Coppola A., Galli C. (1994). Association between hepatitis C and hepatitis E viruses in southern Italy. *Lancet*. **344**: 746-747.

Pogue G. P., Huntley C. C., Hall T. C. (1994). Common replication strategies emerging from the study of diverse groups of positive-strand RNA viruses. *Archives of Virology*. **9** (suppl): 181-194.

Provost P. J., and Hilleman M. R. (1979). Propagation of human hepatitis A in cell culture in vitro. *Proc Soc Exp Biol Med*. **160**(2): 213-221.

Pujol F. H., Favorov M. O., Marcano T., Este J. A., Magris M., Liprandi F., Khudyakov Y. E., Khudyakova N. S., Fields H. A. (1994). Prevalence of antibodies against hepatitis E virus among urban and rural populations in venezuela. *Journal of Medical Virology.* **42**: 234-236.

Purtilo D. T., Hallgren H. M., Yunis E. J. (1972). Depressed maternal lymphocyte response to phytohaemagglutinin in human pregnancy. *Lancet.* **1**: 769-71.

Quiroga J. A., Cotonat T., Castillo I., Carreno V. (1996). Hepatitis E virus seroprevalence in acute viral hepatitis in a developed country confirmed by a supplemental assay. *Journal of Medical Virology.* **50**(1): 16-19.

Rab M. A., Bile M. K., Mubarik M. M., Asghar H., Sami Z., Siddiqi S., Dil A. S., Barzgar M. A., Chaudhry M. A., Burney M. I. (1997). Water-borne hepatitis E virus epidemic in Islamabad, Pakistan: a common source outbreak traced to the malfunction of a modern water treatment plant. *American Journal of Tropical Medicine and Hygiene.* **57**(2): 151-157.

Ranger-Rogez, S., Denis F., Udin L. (1993). Seroprevalence of hepatitis E among pregnant foreign residents in France. *Lancet.* **342**: 998-999.

Rapicetta M., Kondili L. A., Pretolani S., Stroffolini T., Chionne P., Villano U., Madonna E., Casali F., Gasbarrini G. (1999). Seroprevalence and anti-HEV persistence in the general population of the republic of San Marino. *Journal of Medical Virology.* **58**(1): 49-53.

Ray R., Aggrawal R., Salunke P.N., Mehrotra N. N., Talwar G. P., Naik S. R. (1991). Hepatitis E virus genome in stools of hepatitis patients during large epidemic in north India. *Lancet.* **338**: 783-784.

Ray R., Jameel S., Manivel V., Ray R. (1992). Indian hepatitis E virus shows a major deletion in the small open reading frame. *Virology.* **189**: 359-362.

Reichard O., Norkrans G., Fryden A., Braconier J. H., Sonnerborg A., Weiland O. (1998). Randomised, double-blind, placebo-controlled trial of interferon alpha-2b with and without ribavirin for chronic hepatitis C. *Lancet.* **351**: 83-87.

Reyes G. R., Purdy M. A., Kim J. P., Luk K. C., Young L. M., Fry K. E., Bradley and D. W. (1990). Isolation of a cDNA from the virus responsible for enterically transmitted non-A, non-B hepatitis. *Science.* **247**: 1335-1339.

Reyes G. R., and Kim J. P. (1991). Sequence-independent, single-primer amplification (SISPA) of complex DNA populations. *Molecular and Cellular Probes.* **5**: 473-481.

Riddell M. A., Li F., Anderson D. A. (2000). Identification of immunodominant and conformational epitopes in the capsid protein of hepatitis E virus by using monoclonal antibodies. *Journal of Virology.* **74**(17): 8011-8017.

Rizetto M., Canese M. G., Arico S., Crivelli O., Trepo C., Bonino F., Verme G. (1977). Immunofluorescence detection of new antigen-antibody system (delta/anti-delta) associated to hepatitis B in liver and serum of HBsAg carriers. *Gut.* **18**: 997-1003.

Robson S. C., Adams S., Brink N., Woodruff B., Bradley D. (1992). Hospital outbreak of hepatitis E [letter]. *Lancet.* **339**: 1424-1425.

Ropp S. L., Tam A. W., Beames B., Purdy M., Frey T. K. (2000). Expression of the hepatitis E virus ORF 1. *Archives of Virology.* **145**(7): 1321-1337.

Sambrooke J., Fritsch E. F., Maniatis T. (1989). Molecular cloning: A laboratory manual. Second edition. Cold Spring Harbour Laboratory Press, Cold Spring Harbour, N.Y.

Sanger F., Nicklen S., Coulson A. (1977). DNA sequencing with chain-terminating inhibitors. *Proceedings of the National Academy of Sciences.* **74**(12): 5463-5467.

Scharschmidt B. F. (1995). Hepatitis E: a virus in waiting. *Lancet.* **346**: 519-520.

Schiff E. R. (1992). Atypical clinical manifestations of hepatitis A. *Vaccine.* **10** (suppl 1): s18-s20.

Schlauder G. G., Dawson G. J., Mushahwar I. K., Ritter A., Sutherland R., Moaness A., Kamel M. A. (1993). Viraemia in Egyptian children with hepatitis E virus infection [letter]. *Lancet.* **341**: 378.

Schlauder G. G., Pilot-Matias T. J., Gabriel G. S., Simons J. N., Muerhoff A. S., Dawson G. J., Mushahwar I. K. (1995). Origin of GB-hepatitis viruses. *Lancet.* **346**: 447-448.

Schlauder G. G., Dawson G. J., Erker J. C., Kwo P. Y., Knigge M. F., Smalley D. L., Rosenblatt J. E., Desai S. M., Mushahwar I. K. (1998). The sequence and phylogenetic analysis of a novel hepatitis E virus isolated from a patient with acute hepatitis reported in the United States. *Journal of General Virology.* **79**: 447-456. (Genbank accession **AF060668, AF060669**)

Schlauder G. G., Desai S. M., Zanetti A. R., Tassopoulos N. C., Mushahwar I. K. (1999). Novel hepatitis E (HEV) isolates from Europe: Evidence for additional genotypes of HEV. *Journal of Medical Virology.* **57**(3): 243-251.

Schlauder G., Frider B., Sookoian S., Castano G. C., Mushahwar I. K. (2000). Identification of 2 novel isolates of hepatitis E virus in Argentina. *Journal of Infectious Diseases.* **182**: 294-297.

Schlauder G., and Mushahwar I. K. (2001). Genetic heterogeneity of hepatitis E virus. *Journal of Medical Virology.* **65**: 282-292.

Seipp S., Scheidel M., Hofmann W. J., Tox U., Theilmann L., Goeser T., Kallinowski B. (1999). Hepatotropism of GB virus C (GBV-C): GBV-C replication

in human hepatocytes and cells of human hepatoma cell lines. *Journal of Hepatology.* **30**(4): 570-579.

Shidrawi R. G., Skidmore S. J., Coleman J. C., Dayton R., Murray-Lyon I. M. (1994). Hepatitis E - an important cause of imported non-A, non-B hepatitis among migrant workers in Qatar. *Journal of Medical Virology.* **43**: 412-414.

Simmonds P., Alberti A., Alter H. J., Bonino F., Bradley D. W., Brechot C., Brouwer J. T., Chan S. W., Chayama K., Chen D. S. (1994). A proposed system for nomenclature of hepatitis C viral genotypes. *Hepatology.* **19**: 1321-1324.

Simmonds P., Davidson F., Lycett C., Prescott L. E., MacDonald D. M., Ellender J., Yap P. L., Ludlam C. A., Haydon G. H., Gillon J., Jarvis L. M. (1998). Detection of a novel DNA virus (TTV) in blood donors and blood products. *Lancet.* **352**: 191-195.

Simmons D. T., and Strauss J. H. (1972). Replication of Sindbis virus. II. Multiple forms of double-stranded RNA isolated from infected cells. *Journal of Molecular Biology.* **71**: 615-631.

Simons J. N., Pilot-Matias T. J., Leary T. P., Dawson G. J., Desai S. M., Schlauder G. G., Muerhoff S., Erker J. C., Buijk S. L., Chalmers M. L., Van Sant C. L., Mushahwar I. K. (1995). Identification of two flavivirus-like genomes in the GB hepatitis agent. *Proceedings of the National Academy of Science USA.* **92**: 3401-3405.

Simons J. N., Leary T. P., Dawson G. J., Pilot-Matias T. J., Muerhoff A. S., Schlauder G. G., Desai S. M., Mushahwar I. K. (1995b). Isolation of novel virus-like sequences associated with human hepatitis. *Nature Medicine.* **1**(6): 564-569.

Singh V., Singh V., Raje M., Nain C. K., Singh K. (1998). Routes of transmission in the hepatitis E epidemic of Saharanpur. *Tropical Gastroenterology.* **19**(3): 107-109.

Skidmore S. J., Yarbough P. O., Gabor K. A., Tam A. W., Reyes G. R., Flower A. J. (1991). Imported hepatitis E in UK [letter]. *Lancet.* **337**: 1541.

Skidmore S. J., Yarbough P. O., Gabor K. A., Reyes G. R. (1992). Hepatitis E virus: the cause of a waterbourne hepatitis outbreak. *Journal of Medical Virology.* **37**: 58-60.

Soe S., Uchida T., Suzuki K., Komatsu K., Azumi J., Okuda Y., Iida F., Shikata T., Rikihisa T., Mizuno K. (1989). Enterically transmitted non-A, non-B hepatitis in cynomolgus monkeys: morphology and probable mechanism of hepatocellular necrosis. *Liver.* **9**(3): 135-145.

Sreenivasan M. A., Banerjee K., Pandya P. G., Kotak R. R., Pandya P. M., Desai N. J., Vaghela L. H. (1978). Epidemiological investigations of an outbreak of infectious hepatitis in Ahmedabad city during 1975-76. *Indian Journal of Medical Research.* **67**: 197-206.

Sridama V., Pacini F., Yang S. L., Moawad A., Reilly M., DeGroot L. J. (1982). Decreased levels of helper T cells: a possible cause of immunodeficiency in pregnancy. *New England Journal of Medicine.* **307**(6): 352-6.

Stevenson P. (2000). Nepal calls the shots in hepatitis E virus vaccine trial. *Lancet.* **355**: 1623.

Sylvan S. P., Jacobson S. H., Christenson B. (1998). Prevalence of antibodies to hepatitis E virus among hemodialysis patients in Sweden. *Journal of Medical Virology.* **54**(1): 38-43.

Takahashi K., Hoshino H., Ohta Y., Yoshida N., Mishiro S. (1998). Very high prevalence of TT virus (TTV) infection in general population of Japan revealed by a new set of primers. *Hepatology Research.* **12**: 233-239.

Takahashi K., Iwasa Y., Hijikata M., Mishiro S. (2000). Identification of a new human DNA virus (TTV-like mini virus, TLMV) intermediately related to TT virus and chicken anaemia virus. *Archives of Virology.* **145**(5): 979-993.

Tam A. W., Smith M. M., Guerra M. E., Huang C-C., Bradley D. W., Fry K. E., Reyes G. R. (1991). Hepatitis E virus (HEV): Molecular cloning and sequencing of the full length viral genome. *Virology.* **185**: 120-131.
(Genbank accession **M73218**)

Tam A. W., White R., Reed E., Short M., Zhang Y. F., Fuerst T. R., Lanford R. E. (1996). In vitro propagation and production of hepatitis E virus from in vivo-infected primary macaque hepatocytes. *Virology.* **215**: 1-9.

Tan D., Im S. W. K., Yao J. L., Ng M. H. (1995). Acute sporadic hepatitis E infection in southern China. *Journal of Hepatology.* **23**: 239-245.

Tanaka H., Okamoto H., Luengrojanakul P., Chainuvati T., Tsuda F., Tanaka T., Miyakawa Y., Mayumi M. (1998). Infection with an unenveloped DNA virus (TTV) associated with posttransfusion non-A to G hepatitis in hepatitis patients and healthy blood donors in Thailand. *Journal of Medical Virology.* **56**: 234-238.

Tandon B. N., Joshi Y. K., Jain S. K., Gandhi B. M., Mathiesen L. R., Tandon H. D. (1982). An epidemic of non-A, non-B hepatitis in North India. *Indian Journal of Medical Research.* **75**: 739-744.

Tassopoulos N. C., Krawczynski K., Hatzakis A., Katsoulidou A., Delladetsima I., Koutelou M. G., Trichopoulos D. (1994). Role of hepatitis E virus in the etiology of community- acquired non-A, non-B hepatitis in Greece - case report. *Journal of Medical Virology.* **42**: 124-128.

Thomas D. L., Mahley R. W., Badur S., Palaoglu K. E., Quinn T. C. (1993). Epidemiology of hepatitis E virus infection in Turkey. *Lancet* **341** :1561-1562.

Thomas D. L., Yarbough P. O., Vlahov D., Tsarev S. A., Nelson K. E., Saah A. J., Purcell R. H. (1997). Seroreactivity to hepatitis E virus in areas where the disease is not endemic. *Journal of Clinical Microbiology.* **35**: 1244-1247.

Ticehurst J., Popkin T. J., Bryan J. P., Innis B. L., Duncan J. F., Ahmed A., Iqbal M., Malik I., Kapikian A. Z., Legters L. J., Purcell R. H. (1992a). Association of hepatitis E virus with an outbreak of hepatitis in Pakistan: serologic responses and pattern of virus excretion. *Journal of Medical Virology.* **36**: 84-92.

Ticehurst J., Rhodes, Jr L. L., Krawczynski K., Asher L. V., Engler W. F., Mensing T. L., Caudill J. D., Sjogren M. H., Hoke,Jr C. H., LeDuc J. W., Bradley D. W., Binn L. N. (1992b). Infection of owl monkeys *(Aotus trivirgatus)* and cynomolgus monkeys *(Macaca fascicularis)* with hepatitis E virus from Mexico. *Journal of Infectious Diseases.* **165**: 835-845.

Torresi J., Li F., Locarnini S. A., Anderson D. A. (1999). Only the non-glycosylated fraction of hepatitis E virus capsid (open reading frame 2) protein is stable in mammalian cells. *Journal of General Virology.* **80**: 1185-88.

Trautwein C., Kiral G., Tillmann H. L., Witteler H., Michel G., Manns M. P. (1995). Risk factors and prevalence of hepatitis E in German immigrants from the former soviet union. *Journal of Medical Virology.* **45**: 429-434.

Trey C. and Davidson L. S. (1970). The management of fulminant hepatic failure. In: Popper H, Shaffner F, (eds). Progress in liver disease. New York: Grune and Stratton, p282-298.

Tsai J-F., Jeng J-E., Chang W-Y., Lin Z-Y., and Tsai J-H. (1994). Antibodies to hepatitis E virus among Chinese patients with acute hepatitis in Taiwan. *Journal of Medical Virology.* **43**: 341-344.

Tsarev S. A., Emerson S. U., Reyes G. R., Tsareva T. S., Legters L. J., Malik I. A., Purcell R. H. (1992). Characterization of a prototype strain of hepatitis E virus. *Proceedings of the National Academy of Science U. S. A.* **89**: 559-563. (Genbank accession **M80581**)

Tsarev S. A., Emerson S. U., Tsareva T. S., Yarbough P. O., Lewis M., Govindarajan S., Reyes G. R., Shapiro M., Purcell R. H. (1993a). Variation in course of hepatitis E in experimentally infected cynomolgus monkeys. *Journal of Infectious Diseases.* **167**: 1302-1306.

Tsarev S. A., Tsareva T. S., Emerson S. U., Kapikian A. Z., Ticehurst J., London W., Purcell R. H. (1993b). ELISA for antibody to hepatitis E virus (HEV) based on complete open-reading frame-2 protein expressed in insect cells: identification of HEV infection in primates. *Journal of Infectious Diseases.* **168**: 369-378.

Tsarev S. A., Tsareva T. S., Emerson S. U., Govindarajan S., Shapiro M., Gerin J. L., Purcell R. H. (1994a). Successful passive and active immunization of cynomolgus monkeys against hepatitis E. *Proceedings of the National Academy of Science USA.* **91**: 10198-10202.

Tsarev S. A., Tsareva T. S., Emerson S. U., Yarbough P. O., Legters L. J., Moskal T., Purcell R. H.. (1994b). Infectivity titration of a prototype strain of hepatitis E virus in cynomolgus monkeys. *Journal of Medical Virology*. **43**: 135-142.

Tsarev S. A., Tsareva T. S., Emerson S. U., Rippy M. K., Zack P., Shapiro M., and Purcell R. H (1995). Experimental hepatitis E in pregnant rhesus monkeys: failure to transmit hepatitis E virus (HEV) to offspring and evidence of naturally acquired antibodies to HEV. *Journal of Infectious Diseases*. **172**: 31-37.

Tsarev S. A., Tsareva T. S., Emerson S. U., Govindarajan S., Shapiro M., Gerin J. L., Purcell R. H. (1997). Recombinant vaccine against hepatitis E: dose response and protection against heterologous challenge. *Vaccine*. **15**(17-18): 1834-1838.

Tsarev S. A., Binn L. N., Gomatis P. J., Arthur R. R., Monier M. K., van Cuyck-Gandre H., Longer C. F., Innis B. L. (1999). Phylogenetic analysis of hepatitis E virus isolates from Egypt. *Journal of Medical Virology*. **57**: 68-74.

Tsega E., Krawczynski K., Hansson B. G., Nordenfelt E., Negusse Y., Alemu W. (1991). Outbreak of acute hepatitis E virus infection among military personnel in northern Ethiopia. *Journal of Medical Virology*. **34**: 232-236.

Tsega E., Hansson B. G., Krawczynski K., Nordenfelt E. (1992). Acute sporadic viral hepatitis in Ethiopia: causes, risk factors, and effects on pregnancy. *Clinical Infectious Diseases*. **14**: 961-965.

Tucker T. J., Kirsch R. E., Louw S. J., Isaacs S., Kannemeyer J., Robson S. C. (1996). Hepatitis E in South Africa: Evidence for sporadic spread and increased seroprevalence in rural areas. *Journal of Medical Virology*. **50**: 117-119.

Tucker T. J., Smuts H. E., Eedes C., Knobel G. D., Eickhaus P., Robson S.C., Kirsch R. E. (2000). Evidence that GBV-C/hepatitis G virus is primarily a lymphotropic virus. *Journal of Medical Virology*. **61**: 52-58.

Uchida T. (unpublished, Genbank accession **D11093**)

Usmanov R. K., Balayan M. S., Dvoinikova O. V., Alymbayeva D. B., Zamyatina N. A., Kazachkov Y. A., Belov V. I. (1994). Experimental hepatitis-E infection in lambs. *Voprosy Virusologii*. **39**: 165-168.

Van Cuyck-Gandre H., Cockman-Thomas R., Caudill J. D., Asher L. S. V., Armstrong K. L., Hauroeder B., Clements N. J., Binn L. N., Longer C. F. (1998). Experimental african HEV infection in Cynomolgus Macaques (Macaca fascicularis). *Journal of Medical Virology*. **55**: 197-202.

Van de Peer Y., and De Wachter R., (1994). TREECON for windows: a software package for the construction and drawing of evoutionary trees for the Microsoft Windows environment. *Comput. Applic. Biosci.* **10**: 569-570.

Vandenvelde C., Heyvaert F., Maes F., Kadi L., Opstelten R. (1994). Hepatitis E virus infection in Belgian soldiers. *Lancet.* **344**: 747.

Velazquez O., Stetler H. C., Avila C., Ornelas G., Alvarez C., Hadler J. C., Bradley D. W., Sepulveda J. (1990). Epidemic transmission of enterically transmitted non-A, non-B hepatitis in Mexico, 1986-1987. *Journal of the American Medical Association.* **263**: 3281-3285.

Vento S., Garofano T., Renzini C., Cainelli F., Casali F., Ghironzi G., Ferraro T., Concia E. (1998). Fulminant hepatitis associated with hepatitis A virus superinfection in patients with chronic hepatitis C. *New England Journal of Medicine.* **338**: 286-290.

Viswanathan R. (1957). Infectious hepatitis in Delhi (1955-1956): a critical study: epidemiology. *Indian Journal of Medical Research.* **45**(suppl): 1-30.

Von Brunn A., Seebach J., Thyagarajan S. P., Mohanavalli B., Menon T., Froesner G. PCR amplification, cloning and sequence determination of a hepatitis E virus isolate from Madras, India. (unpublished, Genbank accession **X99441**)

Wang C. H., Flehmig B., Moeckli R. (1993). Transmission of hepatitis E virus by transfusion? *Lancet.* **341**: 825-826.

Wang M., Graham D. Y., Estes M. K. (1992). Expression, self-assembly and antigenicity of the Norwalk virus capsid protein. *Journal of Virology.* **66**: 6527-6532.

Wang Y. C., Ling R., Erker J. C., Zhang H. Y., Li H. M., Desai S., Mushahwar I. K., Harrison T. J. (1999). A divergent genotype of hepatitis E virus in Chinese patients with acute hepatitis. *Journal of General Virology.* **80**: 169-177.

Wang Y., Zhang H., Ling R., Li H., Harrison T. J. (2000). The complete sequence of hepatitis E virus genotype 4 reveals an alternative strategy for translation of open reading frames 2 and 3. *Journal of General Virology.* **81**: 1675-1686. (Genbank accession **AJ272108**)

Wei S., Walsh P., Huang R., To S. S. T. (2000). 93G, a novel sporadic strain of hepatitis E virus in South China isolated by cell culture. *Journal of Medical Virology.* **61**: 311-318.

Wong D. C., Purcell R. H., Sreenivasan M. A., Prasad S. R., Pavri K. M. (1980). Epidemic and endemic hepatitis in India: evidence for non-A, non-B virus etiology. *Lancet.* **2**: 876-879.

Worm H. C., Wurzer H., Frosner G. (1998). Sporadic hepatitis E in Austria. *New England Journal of Medicine.* **339**(21): 1554-1555.

Wu J-C., Chen C-M., Chiang T-Y., Sheen I-J., Chen J-Y., Tsai W-H., Huang Y-H., Lee S-D. (2000). Clinical and epidemiological implications of swine hepatitis E virus infection. *Journal of Medical Virology.* **60**: 166-171.

Yarbough P. O., Tam A. W., Fry K. E., Krawczynski K., McCaustland K. A., Bradley, D. W., Reyes G. R. (1991). Hepatitis E virus: identification of type-common epitopes. *Journal of Virology.* **65**: 5790-5797.

Yarbough P. O., Krawczynski K., Tam A. W., McAtee C. P., McCaustland K. A., Zhang Y., Garcon N., Spelbring J., Carson D., Myriam F., Lifson J. D., Slaoui M., Prieels J. P., Margolis H., Fuerst T. R. (1997). Prevention of hepatitis E using r62K subunit vaccine: full protection against heterologous HEV challenge in cynomolgus macaques. (in) Rizzetto M., Purcell R. H., Gerin J. L., Verme G. (Eds); *Viral Hepatitis and Liver Disease*: 650-655; Edizioni Minerva Medica, Torino (1997).

Yin S. R., Purcell R. H., Emerson S. U. A new Chinese isolate of hepatitis E virus: comparison with strains recovered from different geographical regions (unpublished, Genbank accession **L25547**)

Zaaijer H. L., Yin M. F., and Lelie P. N. (1992). Seroprevalence of hepatitis E in The Netherlands. *Lancet.* **340**: 681.

Zaaijer H. L., Kok M., Lelie P. N., Timmerman R. J., Chau K., van der Pal H. J. H. (1993). Hepatitis E in The Netherlands: imported and endemic. *Lancet.* **341**: 826.

Zaaijer H. L., Mauser-Bunschoten E. P., ten-Veen J. H., Kapprell H. P., Kok M., van den Berg H. M., Lelie P. N. (1995). Hepatitis E virus antibodies among patients with hemophilia, blood donors, and hepatitis patients. *Journal of Medical Virology.* **46**: 244-246.

Zafrullah M., Ozdener M. H., Panda S. K., Jameel S. (1997). The ORF3 protein of hepatitis E virus is a phosphoprotein that associates with the cytoskeleton. *Journal of Virology.* **71**(12): 9045-53.

Zafrullah M., Ozdener M. H., Kumar R., Panda S. K., Jameel S. (1999). Mutational analysis of glycosylation, membrane translocation, and cell-surface expression of the hepatitis E virus ORF2 protein. *Journal of Virology.* **73**(5): 4074-4082.

Zanetti A. R., Dawson G. J., Romano L., and the study group of hepatitis E (Tanzi E., Mele A., Rapicetta M., Stroffolini T., Cargnel A., Chircu L. V., Manzillo G., Mazzotta F., Puoti M., Vigano P., Vigevani G. M., Sagliocca L., Bonino P., Dagostino F., Boschini A., Galli C., and Ritter A.). (1994). Hepatitis type E in Italy - a seroepidemiological survey. *Journal of Medical Virology.* **42**: 318-320.

Zanetti A. R., Schlauder G. G., Romano L., Tanzi E., Fabris P., Dawson G. J., Mushahwar I. K. (1999). Identification of a novel variant of hepatitis E virus in Italy. *Journal of Medical Virology.* **57**(4): 356-360.

Zhang Y., McAtee P., Yarbough P. O., Tam A. W., Fuerst T. (1997). Expression, characterization, and immunoreactivities of a soluble hepatitis E virus putative capsid protein species expressed in insect cells. *Clin. Diagn. Lab. Immunol.* **4**(4): 423-428.

Zhang M., Purcell R. H., Emerson S. U. (2001). Identification of the 5' terminal sequence of the SAR-55 and MEX-14 strains of hepatitis E virus and confirmation that the genome is capped. *Journal of Medical Virology.* **65**: 293-295.

Zheng Z. M., He P-J., Caueffield D., Neumann M., Specter S., Baker C. C., Bankowski M. J. (1995). Enterovirus 71 isolated from China is serologically similar to the prototype E71 BrCr strain but differs in the 5'-noncoding region. *Journal of Medical Virology.* **47**: 161-167.

Zhuang H., Cao X. Y., Liu C. B., Wang G. M. (1991). Enterically transmitted non-A, non-B hepatitis in China. In: *Viral hepatitis C, D, E.* Editors: Shikata T., Purcell R. H., and Uchida T., pp 277-285. Elsevier, New York.

Zhuang H., Cao X. Y., Liu C. B., Wang G. M. (1991a). Epidemiology of hepatitis E in China. *Gastroenterologia Japonica.* **26** (suppl 3): 135-138.

APPENDIX 1

Complete genome sequence of Hepatitis E virus associated with acute liver failure (Genbank accession X98292)

1: X98292   Hepatitis E virus, complete genome

```
LOCUS           HEVCG          7202 bp    DNA          VRL        17-JUN-1998
DEFINITION  Hepatitis E virus, complete genome.
ACCESSION   X98292
VERSION     X98292.1  GI:1370097
KEYWORDS    cysteine protease; methyltransferase; RNA-dependent RNA
polymerase;
            structural protein.
SOURCE      Hepatitis E virus.
  ORGANISM  Hepatitis E virus
            Viruses; ssRNA positive-strand viruses, no DNA stage.
REFERENCE   1  (bases 1 to 7202)
  AUTHORS   Donati,M.C., Fagan,E.A. and Harrison,T.J.
  TITLE     Sequence analysis of full length HEV clones derived
            directly from human liver in fulminant hepatitis E
  JOURNAL   (in) Rizzetto,M., Purcell,R.H., Gerin,J.L. and Verme,G
            (Eds.);VIRAL HEPATITIS AND LIVER DISEASE: 313-316;
            Edizioni Minerva Medica, Torino (1997)
REFERENCE   2  (bases 1 to 7202)
  AUTHORS   Harrison,T.J.
  TITLE     Direct Submission
  JOURNAL   Submitted (04-JUN-1996) T.J. Harrison, Royal Free
            Hospital School of Medicine, University Department of
            Medicine, Rowland Hill Street, Hampstead, London NW3 2PF,
            UK
FEATURES             Location/Qualifiers
     source          1..7202
                     /organism="Hepatitis E virus"
                     /isolate="hev037"
                     /specific_host="Homo sapiens"
                     /db_xref="taxon:12461"
                     /note="isolated from Indian liver; acute liver
                     failure"
     gene            27..5108
                     /gene="ORF1"
     CDS             27..5108
                     /gene="ORF1"
                     /note="RNA helicase; RNA-dependent RNA
                      polymerase; methyl transferase"
                     /codon_start=1
                     /product="cysteine protease"
                     /protein_id="CAA66936.1"
                     /db_xref="GI:1370098"
                     /db_xref="SPTREMBL:Q69410"
```

```
/translation="MEAHQFIKAPGITTAIEQAALAAANSALANAVVVRPFLSHQQIE
ILINLMQPRQLVFRPEVFWNHPIQRVIHNELELYCRARSGRCLEIGAHPRSINDNPNV
VHRCFLRPVGRDVQRWYTAPTRGPAANCRRSALRGLSAADRTYCFDGFSGCNFPAETG
IALYSLHDMSPSDVAEAMFRHGMTRLYAALHLPPEVLLPPGTYRTASYLLIHDGRRVV
VTYEGDTSAGYNHDVSNLRSWIRTTKVTGDHPLVIERVRAIGCHFVLLLTAAPEPSPM
PYVPYPRSTEVYVRSIFGPGGTPSLFPTSCSTKSTFHAVPAHIWDRLMLFGATLDDQA
FCCSRLMTYLRGISYKVTVGTLVANEGWNASEDALTAVITAAYLTICHQRYLRTQAIS
KGMRRLEREHAQKFITRLYSWLFEKSGRDYIPGRQLEFYAQCRRWLSAGFHLDPRVLV
```

FDESAPCHCRTAIRKAVSKFCCFMKWLGQECTCFLQPAEGAAGDQGHDNEAYEGSDVD
PAESAISDISGSYVVPGTALQPLYQALDLPAEIVARAGRLTATVKVSQVDGRIDCETL
LGNKTFRTSFVDGAVLETNGPERHNLSFDASQSTMAAGPFSLTYAASAAGLEVRYVAA
GLDHRAVFAPGVSPRSAPGEVTAFCSALYRFNREAQRHALTGNFWFHPEGLLGLFAPF
SPGHVWESANPFCGESTLYTRTWSEVDAVSSPARPDLGFASEPSIPSRAATPTPAALQ
PSSAPDPFPPPSAPALGEPAPGVTAVAPAITHQTARHRRLLFTYPDGSKVFAGSLFES
TCTWLVNASNVDHRPGGGLCHAFYQRYPTSFDAASFVMRDGAAAYTLTPRPIIHAVAP
DYRLEHNPKRLEAAYRETCSRLGTAAYPLLGTGIYQVPIGPSFDAWERNHRPGDELYL
PELAARWFEANRPACPTLTITEDAARTANLAIELDSATDVGRACAGCRVTPGVVQYQF
TAGVPGSGKSRSITQADVDVVVVPTRELRNAWRRRGFAAFTPHTAARVTQGRRVVIDE
APSLPPHLLLLHMQRAATVHLLGDPNQIPAIDFEHAGLVPAIRPDLAPTSWWHVTHRC
PADVCELIRGAYPMIQTTSRVLRSLFWGEPAVGQKLVFTQAAKAANPGSVTVHEAQGA
TYTETTIIATADARGLIQSSRAHAIVALTRHTEKCVIIDAPGLLREVGISDAIVNNFF
LAGGEIGHQRPSVIPRGNPDTNVDTLAAFPPSCQISAFHQLAEELGHRPAPVAAVLPP
CPELEQGLLYLPQELTTCDSVVTFELTDIVHCRMAAPSQRKAVLSTLVGRYGRRTKLY
NASHSDVRDSLARFIPTIGPVQVTTCELYELVEAMVEKGQDGSAVLELDLCNRDVSRI
TFFQKDCNKFTTGETIAHGKVGQGISAWSKTFCALFGPWFRAIEKAILALLPQGVFYG
DAFDDTVFSAAVAAAKASMVFENDFSEFDSTQNNFSLGLECAIMEECGMPQWLIRLYH
LIRSAWILQAPKESLRGFWKKHSGEPGTLLWNTVWNMAVITHCYDFRDLQVAAFKGDD
SIVLCSEYRQSPGAAVLIAGCGLKLKVDFRPIGLYAGVVVAPGLGALPDVVRFAGRLT
EKNWGPGPERAEQLRLAVSDFLRKLTNVAQMCVDVVSRVYGVSPGLVHNLIGMLQTVA
DGKAHFTESVKPVLDLTNSILCRVE"
       gene   5105..5476
                     /gene="ORF3"
       CDS   5105..5476
                     /gene="ORF3"
                     /codon_start=1
                     /protein_id="CAA66938.1"
                     /db_xref="GI:1370100"
                     /db_xref="SPTREMBL:Q69412"

/translation="MNNMSFAAPMGSRPCALGLFCCCSSCFCLCCPRHRPVSRLAAVV
GGAAAVPAVVSGVTGLILSPSQSPIFIQPTPLPPMSPLRPGLDLAFANPPDHSAPLGV
TRPSAPPLPHVVDLPQLGPRR"
       gene   5146..7128
                     /gene="ORF2"
       CDS   5146..7128
                     /gene="ORF2"
                     /codon_start=1
                     /product="structural protein"
                     /protein_id="CAA66937.1"
                     /db_xref="GI:1370099"
                     /db_xref="SPTREMBL:Q69411"

/translation="MRPRPILLLLLMFLPMLPAPPPGQPSGRRRGRRSGGSGGGFWGD
RVDSQPFAIPHIHPTNPFAPDVTAAAGAGPRVRQPARPLGSAWRDQAQRPAATSRRRP
TTAGAAPLTAVAPAHDTPPVPDVDSRGAILRRQYNLSTSPLTSPVATGTNLVLYAAPL
SPLLPLQDGTNTHIMATEASNYAQYRVARATIRYRPLVPNAVGGYAISISFWPQTTTT
PTSVDMNSITSTDVRILVQPGIASELVIPSERLHYRNQGWRSVETSGVAEEEATSGLV
MLCIHGLPVNSYTNTPYTGALGLLDFALEFEFRNLTPGNTNTRVSRYSSTARHRLRRG
ADGTAELTTTAATRFMKDLYFTSTNGVGEIGRGIALTLFNLADTLLGGLPTELISSAG
GQLFYSRPVVSANGEPTVKLYTSVENAQQDKGIAIPHDIDLGESRVVIQDYDNQHEQD
RPTPSPAPSRPFSVLRANDVLWLSLTAAEYDQSTYGSSTGPVYVSDSVTLVNVATGAQ
AVARSLDWTKVTLDGRPLSTIQQYSKTFFVLPLRGKLSFWEAGTTKAGYPYNYNTTAS
DQLLIENAAGHRVAISTYTTSLGAGPVAISAVAVLAPHSALALLEDTMDYPARAHTFD
DFCPECRPLGLQGCAFQSTVAELQRLKMKVGKTREL"

```
BASE COUNT      1238 a    2284 c    1861 g    1819 t
ORIGIN
1     ggcagaccac ctatgtggtc gatgccatgg aggcccatca gtttatcaag gctcctggca
61    tcactactgc tattgagcag gctgctctgg cagcggccaa ctctgccctg gcgaatgctg
121   tggtagttag gccctttctc tctcaccagc agattgagat cctcattaac ctaatgcaac
181   ctcgccagct tgttttccgc cccgaggttt tctggaacca ccccatccag cgtgttatcc
241   acaatgagtt ggagctttac tgccgtgccc gctccggccg ctgtcttgag attggtgccc
301   atccccgctc aataaatgac aatcctaatg tggtccaccg ctgcttcctc cgccctgtcg
361   ggcgcgatgt tcagcgctgg tatactgccc ccactcgcgg gccagctgct aattgccggc
421   gttccgcact gcgcgggctc tctgccgctg accgcactta ctgcttcgac gggttttctg
481   gttgtaactt ccccgccgaa acgggcatcg ccctctattc cctccatgat atgtcaccat
541   ctgatgtcgc cgaggccatg ttccgccatg gtatgacgcg gctttatgcc gccctccatc
601   ttccgcctga ggtcttgttg cctcctggca cataccgcac cgcgtcgtat ctgctgatcc
661   atgacggcag gcgcgttgtg gtgacgtacg agggtgacac tagtgctggt tataaccatg
721   atgtttccaa cctgcgctcc tggattagaa ccactaaggt caccggagac caccctctcg
781   ttatcgagcg ggttagggcc attggctgtc actttgttct cttactcacg gctgccccgg
841   agccatcacc tatgccttat gtcccttacc cccggtctac cgaggtctat gtccgatcga
901   tcttcggccc gggcggcacc ccttccctat tcccaacctc atgctccact aagtcgacct
961   tccatgccgt ccctgcccat atttgggacc gtcttatgct gttcggggcc accctggatg
1021  accaagcctt ttgctgctcc cgtctaatga cctaccttcg cggcattagc tacaaggtta
1081  ctgtgggcac tcttgtggcc aatgaaggct ggaacgcctc tgaggacgct ctcacagctg
1141  tcattactgc cgcctacctt accatctgcc accagcggta tctccgcacc caagctatat
1201  ctaagggggat gcgtcgtttg gagcgggagc atgcccagaa attcataaca cgcctctaca
1261  gttggctttt tgagaagtcc ggccgtgatt atatccccgg ccgtcagttg gagttctacg
1321  cccagtgtag gcgctggctc tcggccggct ttcaccttga cccacgggtg ctggttttg
1381  atgagtcggc ccctgccac tgcaggaccg cgattcgtaa ggcggtctca agttttgct
1441  gctttatgaa gtggcttggc caggaatgta cctgctttct acagcctgca gagggtgccg
1501  ctggcgacca gggccatgat aatgaagcct atgagggggtc tgatgttgac cctgctgaat
1561  ccgccattag tgatatatct gggtcctacg tagtccctgg cactgccctc cagccgcttt
1621  atcaggccct tgacctcccc gctgagattg tggcccgtgc aggccggctg accgccacag
1681  taaaggtctc ccaggtcgac gggcggatcg attgtgagac ccttcttggt aacaaaacct
1741  ttcgcacgtc gttcgttgac ggggcggtct tagagaccaa tggcccagag cgccacaatc
1801  tctcttttga tgccagtcag agcactatgg ccgctggccc cttcagtctt acctatgccg
1861  cctctgccgc tgggctggag gtgcgctatg ttgctgctgg gcttgaccac cgggcggttt
1921  tcgcccccgg cgtttcaccc cggtcggccc ctggcgaggt caccgccttt tgctctgccc
1981  tatacaggtt taatcgtgaa gcccagcgtc atgcgctgac cggtaatttt tggttccatc
2041  ctgaggggct ccttggcctc tttgccccgt tttcgcccgg gcatgtctgg gagtcggcta
2101  atccattttg tggcgagagc acactctaca cccgtacttg gtcggaggtt gatgctgttt
2161  ctagtccagc ccggcctgac ttaggttttg catctgagcc ttctatacct agcagggccg
2221  ccacacctac cccggcggct cttcaaccct cctctgcacc ggaccctttc cccccccct
2281  ctgctccggc gcttggtgag ccggcccctg gcgttactgc cgtggccccg gccataaccc
2341  accagacggc tcggcatcgc cgcctgctct tcacctaccc ggatggctcc aaggtatttg
2401  ccggctcgct gtttgagtcg acatgtacgt ggcttgttaa cgcgtctaat gttgaccatc
2461  gccctggcgg cgggctctgc catgcatttt accaaaggta ccccaccctcc tttgatgctg
2521  cctctttgt gatgcgcgac ggcgcggccg cttatacatt aaccccccgg ccaataattc
2581  atgccgtcgc ccctgattat aggttggaac ataacccaaa gaggcttgag gccgcctacc
2641  gggagacttg ctcccgtctc ggcaccgctg catacccgct cctcgggacc ggcatatacc
2701  aggtaccgat cggtcccagt ttcgacgcct gggagcggaa ccaccgcccc ggtgatgagt
2761  tgtatcttcc tgagcttgct gccagatggt tcgaggccaa taggccggcc tgcccaactc
2821  ttactataac tgaggatgct gcacggacag cgaacctggc catcgagctt gactcagcca
2881  cagatgtcgg ccgggcttgt gctggttgtc gagtcacccc cggcgttgtg cagtaccagt
2941  ttactgcagg tgtgcctggg tccggcaagt cccgctctat cacccaagcc gacgtggacg
3001  ttgtcgtggt cccgacccgg gagttgcgta atgcctggcg ccgccgcggc ttcgctgctt
3061  tcaccccgca cactgctgct agagtcactc aggggcgccg ggttgtcatt gatgaagctc
3121  catctcttcc cccccacctg ctgctgctcc acatgcagcg ggccgccacc gttcaccttc
3181  ttggcgaccc gaaccagatc ccagccatcg actttgagca cgccgggctc gtccccgcca
3241  tcaggcctga cctagcccct acctcctggt ggcatgttac ccatcgttgc cctgcggatg
3301  tatgtgagct gatccgtggt gcataccta tgatccagac cactagccgg gtcctccgtt
3361  cgttgttctg gggtgagcct gccgttgggc agaaactagt gttcacccag gcggctaagg
3421  ctgccaaccc cggttcagtg acggtccacg aggcacaggg cgctacctat acagagacca
3481  ccattatcgc tacggcagat gcccgaggcc tcattcagtc gtctcgggct catgccattg
3541  ttgccctgac gcgccatact gagaagtgcag tcatcattga cgcgcaggac ctgcttcgcg
3601  aggtgggcat ttccgatgca atcgttaata actttttcct tgccggtggc gaaattggcc
3661  accaacgccc gtctgttatc cctcgcggca accctgacac aaacgttgac accttggctg
3721  ccttcccgcc gtcttgccag attagcgcct ccaccagtt ggctgaggag cttggccata
3781  gacctgcccc tgttgcagct gttctaccgc cctgccctga gctcgaacag ggcctcctct
3841  atctgcccca ggagctcacc acctgtgata gtgttgtaac atttgaatta acagatattg
```
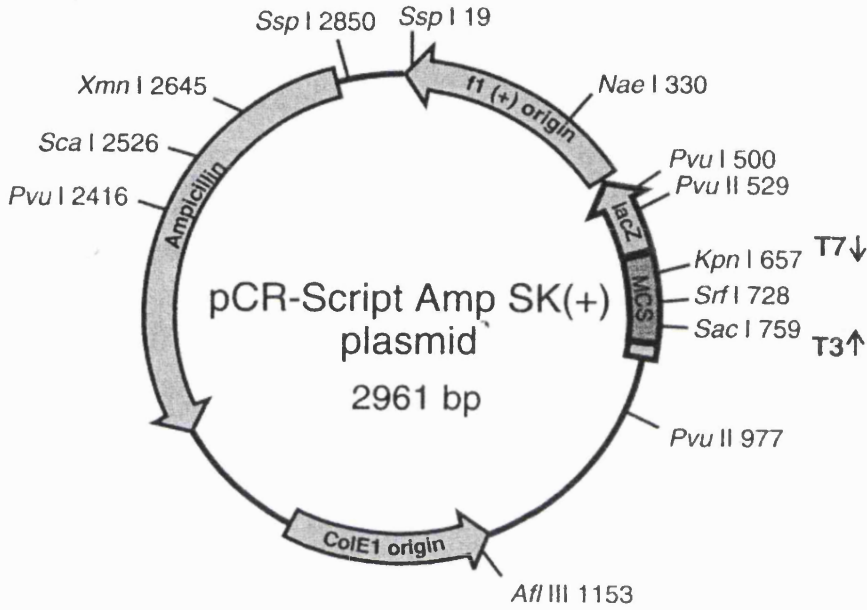
192

```
3901 tgcattgtcg catggccgcc ccgagtcagc gcaaggccgt gctgtccacg ctcgtgggcc
3961 gttacggccg tcgcacaaag ctctacaatg cctctcactc tgatgttcgt gactctcttg
4021 cccgttttat cccgaccatt ggccccgtac aagttacaac ctgtgaattg tatgagctag
4081 tggaggccat ggtcgagaag ggccaggatg gctccgccgt ccttgagctc gacctctgca
4141 accgtgacgt atccaggatc accttcttcc agaaagattg taataagttt accacgggag
4201 agaccatcgc ccatggtaaa gtgggccagg gcatttcggc ctggagtaag accttttgtg
4261 cccttttttgg cccctggttc cgcgctattg agaaggctat tctggcccta ctccctcagg
4321 gtgtgtttta tggagatgcc tttgacgaca ctgtcttctc ggcggctgtg gccgcagcaa
4381 aggcgtccat ggtgtttgag aatgactttt ccgagtttga ctccacccag aataactttt
4441 ccttgggcct agagtgcgct atcatggagg agtgcgggat gccgcagtgg ctcatccgct
4501 tgtaccacct tataaggtct gcgtggatcc tgcaggcccc gaaggagtct ctgcgagggt
4561 tttggaagaa acactccggt gagcccggca ctcttctttg gaatactgtc tggaacatgg
4621 ccgttattac ccactgttat gattttcgtg atttgcaggt ggctgccttt aaaggtgatg
4681 attcgatagt gctttgcagt gagtaccgtc agagtccagg agctgctgtc ctgattgctg
4741 gttgtggctt aaagctgaag gtggattttc gccctattgg tctgtatgca ggtgttgtgg
4801 tggccccccgg ccttggcgcg cttcctgacg ttgtgcgctt cgccggccgg ctcactgaga
4861 agaattgggg ccctggcccc gagcgggcgg agcagctccg ccttgctgta agtgatttcc
4921 tccgcaagct tacgaatgta gctcagatgt gtgtggatgt tgtttctcgt gtttatgggg
4981 tttctcctgg gctcgttcat aacctgattg gcatgctaca gactgttgct gatggcaagg
5041 cccatttcac tgagtcagta aagccagtgc tcgacctgac aaattcaatc ctgtgtcggg
5101 tggaatgaat aacatgtctt ttgctgcgcc catgggttcg cgaccatgcg ccctcggcct
5161 attctgttgt tgctcctcat gtttttgcct atgttgcccg cgccaccgcc cggtcagccg
5221 tctggccgcc gtcgtgggcg gcgcagcggc ggttccggcg gtggtttctg gggtgaccgg
5281 gttgattctc agcccttcgc aatcccccat attcatccaa ccaacccctt tgccccccgat
5341 gtcaccgctg cggccggggc tggacctcgc gttcgccaac ccgcccgacc actcggctcc
5401 gcttggcgtg accaggccca gcgccccgcc gctacctcac gtcgtagacc taccacagct
5461 ggggccgcgc cgctaaccgc ggtcgccccg gctcatgaca ccccgccagt gcctgatgtc
5521 gactcccgcg gcgctatctt gcgccggcag tataacctat caacatcccc tctcacctct
5581 cccgtggcca ccggcacaaa tttggtcctt tatgccgccc ctcttagccc gcttttaccc
5641 cttcaggacg gcaccaatac tcatataatg gctacagaag cctctaacta tgcccagtac
5701 cgggttgccc gtgccacaat tcgttatcgc ccgcttgtcc ccaatgctgt tggtggttat
5761 gccatctcca tctcgttctg gccacaaacc accaccactc cgacgtccgt tgatatgaat
5821 tcaataactt cgacggatgt tcgtattttta gttcagcccg gcatagcctc tgagcttgtg
5881 atcccaagtg agcgcttaca ctatcgtaac caagggtggc gctctgttga gacctccggg
5941 gtggcggaag aggaggccac ctctggtctc gttatgcttt gcatacatgg cttacccgta
6001 aattcttaca ctaatacacc ctacaccggc gccctcgggc tgttggattt tgcccttgaa
6061 tttgagttcc gcaacctcac tcccggcaac accaatactc gggtctctcg ttactccagc
6121 actgctcgcc accgtcttcg ccgcggtgca gatgggaccg ccgagctcac caccacggct
6181 gccacccgct ttatgaagga tctctatttt actagtacta atggtgtcgg tgagatcggc
6241 cgcgggatag ccctaaccct gtttaatctt gctgacactc tgcttggcgg cctaccgaca
6301 gaattgattt cgtcggctgg tggccagctg ttctactccc gtcccgtcgt ctcagccaat
6361 ggcgagccga ctgttaagct gtatacatct gtagagaatg ctcagcagga taagggtatt
6421 gcaatcccgc atgacatcga cctcggggaa tctcgtgtgg tcattcagga ttatgataac
6481 caacatgagc aagatcggcc gacaccttct ccagccccat cgcgcccttt ttctgtcctc
6541 cgagctaatg atgtgctttg gctttctctc accgctgccg agtatgacca gtccacttac
6601 ggttcttcga ccggcccagt ctatgtttct gactctgtga ccctggttaa tgttgcgacc
6661 ggcgcgcagg ccgttgcccg gtcgcttgac tggaccaagg tcacacttga tggtcgcccc
6721 ctctccacca tccagcagta ttcaaagacc ttcttcgtcc tgccgctccg cggtaagctc
6781 tccttctggg aggcaggtac cactaaggcc gggtacccct ataattataa taccactgct
6841 agcgatcaac tgcttattga gaatgccgcc gggcatcggg ttgctatttc cacttacacc
6901 actagcctgg gtgctggccc cgttgccatc tctgcggttg ctgtttttggc cccccactct
6961 gcgctagccc tgcttgagga taccatggac taccctgccc gcgcccacac ttttgatgat
7021 ttctgcccag agtgccgccc ccttggcctc cagggctgtg ctttccagtc tactgtcgct
7081 gagcttcagc gccttaagat gaaggtgggt aaaactcggg agttatagtt tatttgcttg
7141 tgccccccctt ctttctgctg cttatttctc ttttctgcgt ttcgcgctcc ctgaaaaaaa
7201 aa
```

193

APPENDIX 2

Plasmids used in cloning: pCR-Script (Stratagene Inc., USA)

pCR-Script Amp SK(+) plasmid
2961 bp

Ssp I 2850   *Ssp* I 19   *Nae* I 330
Xmn I 2645   f1 (+) origin
Sca I 2526   *Pvu* I 500
Pvu I 2416   *Pvu* II 529
Ampicillin   lacZ
Kpn I 657 T7↓
Srf I 728
Sac I 759 T3↑
MCS
Pvu II 977
ColE1 origin
Afl III 1153

Plasmids used in cloning: pBluescript (Stratagene Inc., USA)
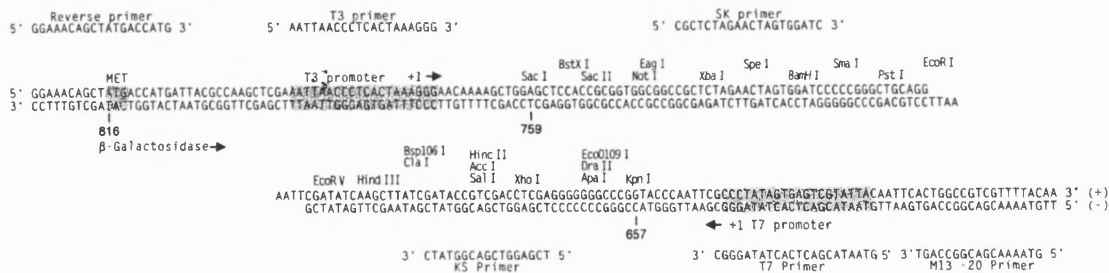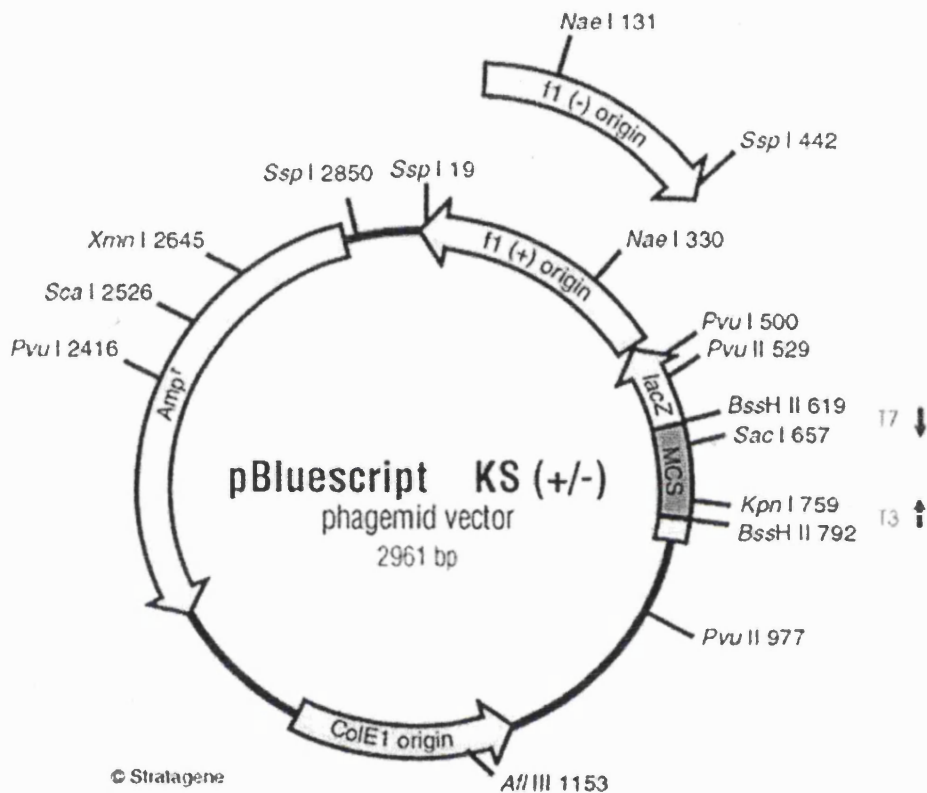
APPENDIX 4

Plasmids used in cloning: pCR2.1 (Invitrogen BV, The Netherlands)
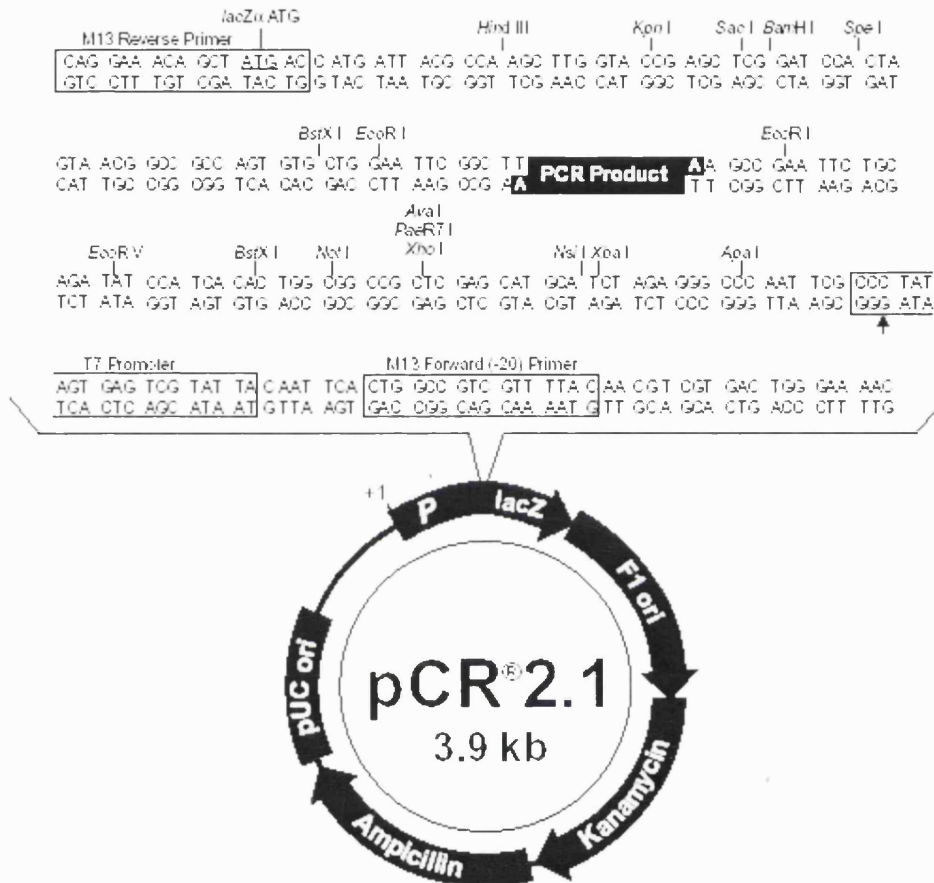
The pCRII vector was identical except for an additional Sp6 promoter sequence and *Nsi*I site located between the M13 reverse primer site and the *Hind*III site.

Plasmids used in cloning: pCI-neo (Promega Inc., USA)

T7 Transcription Start

5'. . . TTAATACGACTCACTATAGGCTAGCCTCGAGAATTCACGCGTGGT

T7 Promoter   Nhe I   Xho I   EcoR I   Mlu I

T3 Transcription Start

ACCTCTAGAGTCGACCCGGGCGGCCGCTTCCCTTTAGTGAGGGTTAATG . . . 3'

Xba I   Acc I   Sma I   Not I   T3 Promoter
          Sal I
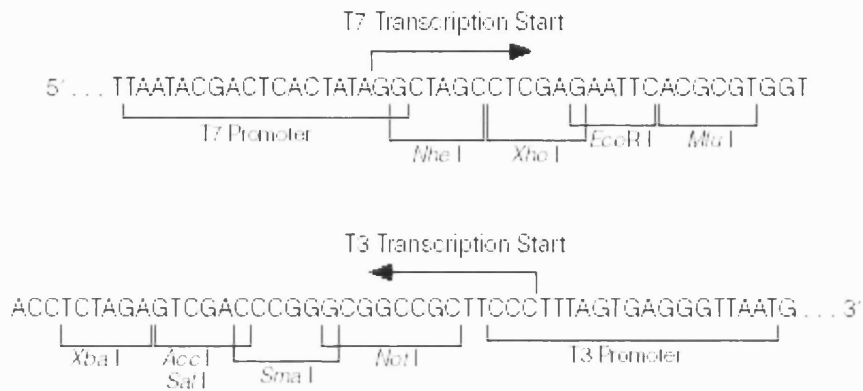
Figure 1. T7 and T3 RNA polymerase promoters and multiple cloning site sequence of the pCI-neo Vector. The sequence shown corresponds to RNA synthesized by T7 RNA polymerase and is complementary to RNA synthesized by T3 RNA polymerase. The strand shown is the same as the ssDNA produced by the vector.

T7 ↓	
Nhe I	1085
Xho I	1091
EcoR I	1096
Mlu I	1102
Xba I	1114
Sal I	1120
Acc I	1121
Sma I	1127
Not I	1131
T3 ↑	

pCI-neo Vector (5472bp)

Bgl II, ori, CMV I.E. Enhancer/Promoter, Sgf I, I-Ppo I, Intron, Amp', SV40 Late poly(A), f1 ori, Synthetic poly(A), SV40 Enhancer/Early Promoter, neo, BamH I
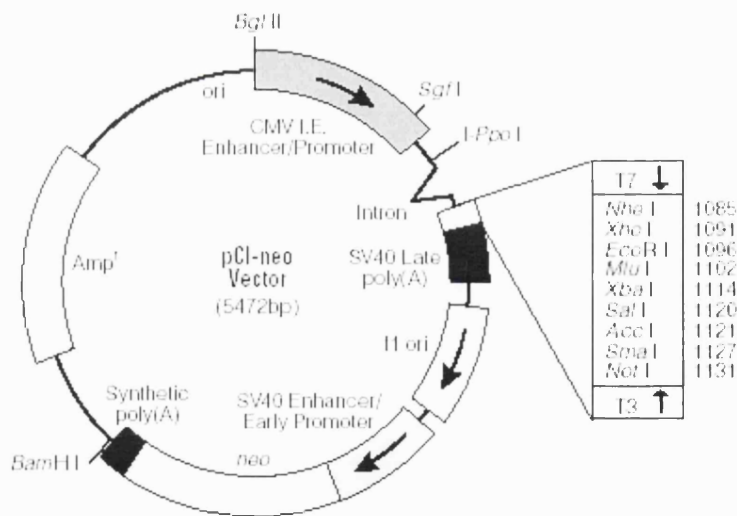
APPENDIX 6

Plasmids used in cloning: pCNTR (5 prime 3 prime Inc., USA)
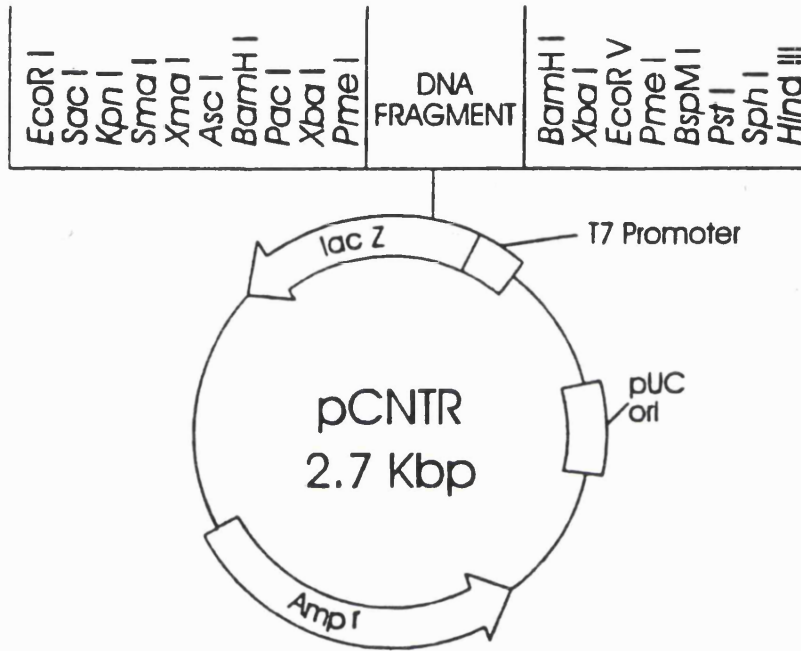
pCNTR Polylinker and Flanking Sequence
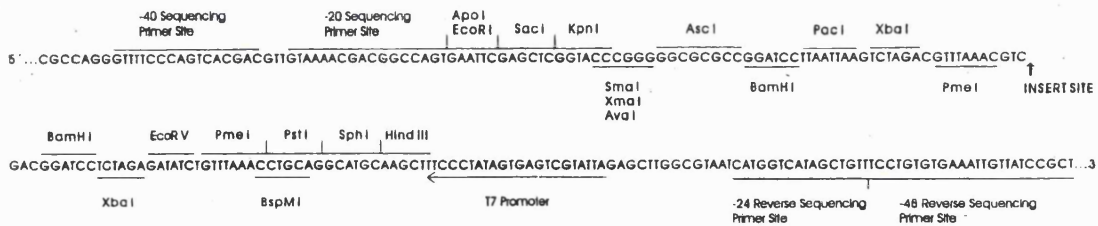
APPENDIX 7

Linear maps of full length HEV cDNA plasmid clones

The linear maps highlight the type of plasmid vector used along with the relative positions of the vector multiple cloning sites, the orientation of the full length HEV sequence and the *Eco*RI and *Hin*dIII restriction endonuclease sites within it.

Vector: pCRII (3.9 kb)

Clones **B2, B55, A3**

*Hin*dIII, *Bam*HI, *Eco*RI   *Hin*dIII   *Eco*RI   *Eco*RI, *Not*I, *Apa*I

           1   4926   5818   7193

Clones **A20, A9**

*Hin*dIII, *Bam*HI, *Eco*RI   *Eco*RI, *Not*I, *Apa*I

                 *Eco*RI   *Hin*dIII

       7193   5818   4926   1
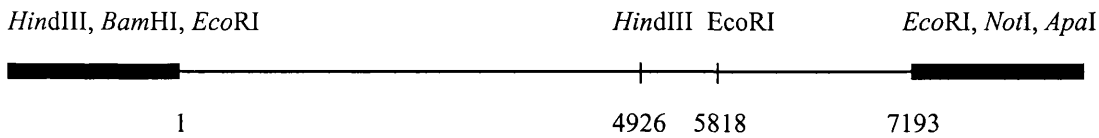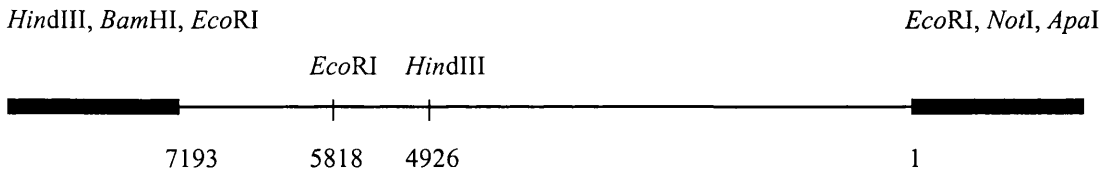
Vector pCNTR (2.7 kb)

Clone **GC5**

*Eco*RI, *Bam*HI, *Pme*I   *Bam*HI, *Pme*I, *Pst*I

                 *Eco*RI   *Hin*dIII

       7193   5818   4926   1
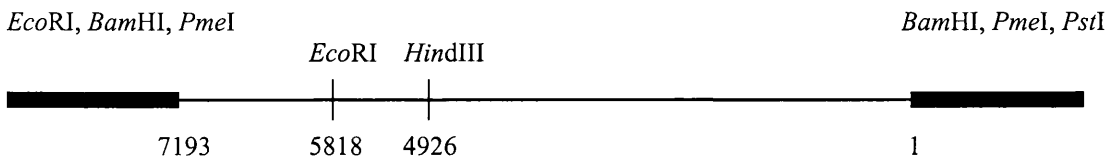
Key: Heavy type line is vector sequence, fine line type is HEV sequence.

APPENDIX 8

Map of restriction enzyme sites used in subcloning

```
                                                    HincII                        PstI
                                                    955                          1404
+---+---+---+---+---+---+---+---+---+---+---+---+---+---+
0   100 200 300 400 500                          1000                        1400

PstI  HincII    HincII                                HincII 2453
1490  1547      1697 1757                            2419 2438
+---+---+---+---+---+---+---+---+---+---+---+---+---+
1400          1700                2000              2300          2600        2800

        PstI  SacII                           HincII
        2948  3047                            3707
+---+---+---+---+---+---+---+---+---+---+---+---+---+
2800          3100              3400          3700          4000        4200


            BamHI PstI                  HindIII                    SacII  HincII SacII
            4525 4534                   4926                       5480   5520  5529
+---+---+---+---+---+---+---+---+---+---+---+---+---+
4200          4500              4800          5100          5400              5600
                                                                            5529
                                                                            SacII
            EcoRI               SacII  SacII                 SacII
            5817                6144   6243                  6771
+---+---+---+---+---+---+---+---+---+---+---+---+---+
5600          5900              6200          6500          6800        7000



+---+---+
7000    7200
```

Restriction endonuclease sites shown are those used in orientation checking of clones and sub-cloning for sequencing purposes and relate to the consensus sequence for the study isolate.

Restriction enzyme names with the cutting position are given above the line; the scale numbering is given below the line.

APPENDIX 9

Basic Molecular Biology Techniques

1. Restriction endonuclease digestion

Typically the following were mixed in a 1.5ml microcentrifuge tube:
DNA
Enzyme (use excess, approx 5u/µg DNA)
Buffer supplied with enzyme
Water up to smallest volume that results in enzyme occupying no more than 10% of total reaction volume.
Incubated for a minimum of 1 hour at the optimum temperature for the enzyme (30°C SmaI, 37°C others).

2. Ligation of DNA

A standard ligation comprised:
Digested vector (approx 100ng)
DNA to clone to give a molar ratio of insert to vector of between 1:1 and 5:1
Ligation buffer supplied with enzyme
T4 bacteriophage DNA ligase 1 Weiss unit
The reaction was allowed to procede usually at 14°C overnight, although room temperature for several hours was sometimes used.

3. Dephosphorylation of DNA

One of two phosphatases was used where the removal of 5' phosphates from ds DNA was required for efficient ligations. When Calf Intestinal Phosphatase was used the enzyme had to be removed by purifying the DNA. The procedure was simplified by using Shrimp Alkaline Phosphatase (USB inc.,USA) which retains activity in most restriction endonuclease buffers and can be inactivated by heating to 65°C for 15 min. The amount of enzyme used depended on the kind of termini and the amount of DNA: only blunt ends were treated in reactions using at least 0.2 units of enzyme per 1.0 pmol termini (2.5µg of 3 Kb plasmid) left at 37°C for 1 hour.

4. Extraction of nucleic acid by Phenol:Chloroform

Deproteinization of nucleic acid solutions was routinely done by extraction first with phenol:chloroform and then with chloroform. The use of phenol:chloroform:isoamyl alcohol (25:24:1) prevented the dissolution of RNAs with long tracts of polyA in the phenol and helped to further inactivate RNAses.
Extraction was done by adding an equal volume of phenol:chloroform:isoamyl alcohol to the nucleic acid in a 1.5 ml polypropylene microcentrifuge tube and vigorously mixing to form an emulsion. The phases were separated by centrifugation at 12000g for 2 min. The nucleic acid partitioned to the upper aqueous phase providing the phenol was equilibrated to pH 7.8-8.0; if the aqueous phase was dense due to >0.5 M salt or >10% sucrose then it formed the lower phase. The aqueous phase was transferred to a clean tube and an equal volume of chloroform added,

mixed, and the phases separated as before. The nucleic acid was recovered by precipitation with ethanol. Back extraction of the organic phase and serial extractions were usually unnecessary.

5. Ethanol precipitation of nucleic acid

Precipitation of nucleic acid was typically done in 70% ethanol, 0.3M sodium acetate pH5.2. Ice cold ethanol was mixed with the nucleic acid solution along with the sodium acetate and centrifuged at 12000g for 15 min at 4°C in a microcentrifuge. Pellets were then washed with 1ml of ice cold 70% ethanol, to remove the salt, and re-centrifuged as above for 10 min. After removal of the supernatant the nucleic acid was air dried and redissolved in buffer / water. When very small fragments (< 100bp) or small quantities (< 20ng / ml) had to be recovered centrifugation was for longer.

6. Agarose gel electrophoresis

1xTAE or 1xTBE buffer (500ml for mini-gel, 1000ml for midi-gel) containing 0.5µg/ml ethidium bromide (1 in 20,000 dilution of 10mg/ml stock) was used.
This buffer was used to make the agarose gel of between 1-2% w/v by melting in a microwave (30mls mini, 80mls midi). The gel was allowed to cool to hand hot and poured to a thickness of 6mm. Electrophoresis was performed in the appropriate buffer at between 25-100V.

7. LMP agarose DNA separation/extraction

DNA fragments were separated on a 1% LMP agarose 1xTAE gel. The DNA band was cut out using a clean scalpel blade and placed in 1.5ml microcentrifuge tube. 1X TAE was added to a total volume of 600µl, heated to 70°C until agarose melted then placed at 37°C. The DNA was extracted with room temperature buffered phenol by vortexing and spinning in a microcentrifuge for 10 mins. The upper aqueous phase was placed in a clean tube and extracted with phenol/chloroform. The DNA was ethanol precipitated from the upper aqueous phase, the pellet washed in 75% ethanol, air dried and redissolved in water or TE (pH8.0).

8. Competent cells

10mls of YT or LB was seeded with a single colony of TG2 or XL2 MRF' *E.coli* and shaken at 225 rpm, 37°C overnight in a rotary incubator. 100ml of LB or TY was inoculated with 1ml of the overnight culture and shaken at 225 rpm, 37°C until at an optical density of 0.2 at 600nm was achieved. Two lots of 45 mls was aliquoted into ice cold sterile Falcon tubes and cooled on ice for 20 mins prior to centrifugation at 4200rpm for 6 min (MSE Mistral 2000, UK). Supernatants were poured off and each pellet resuspended in 15 ml ice cold sterile 0.1M $CaCl_2$. Cells were re-centrifuged as above. Both pellets were resuspended in a total of 10 ml fresh ice cold 0.1M $CaCl_2$ and transferred to a sterile Universal prior to centrifuging at 4200rpm for 5 min. The cells were finally resuspended in 1.5-2.0 ml fresh ice cold 0.1M $CaCl_2$. 100µl aliquots were used. Average efficiencies were around 10^5cfu/µg DNA.

9.   Transformation of competent bacteria

Typically around 10ng of plasmid equivalent was added to an aliquot of competent cells in a Falcon 2059 tube and swirled to mix. This was left on ice for 30 min then heat shocked at 42°C for 45 sec and returned to ice for 2 min. 400µl of pre-warmed SOC (appendix 10) was added and the mixture incubated at 37°C for 1 hour shaking at 225rpm. Initially, 100µl of this was plated onto appropriate selection media. The remainder was stored at 4°C and replated as necessary.

10.   Mammalian cell culture
10.1   Maintenance of cells

Aseptic technique was used in a class 2 cabinet throughout. Cells to be maintained were generally grown in 175 cm² flasks containing sufficient medium to cover the monolayer by several mm. Medium was DMEM containing 10% foetal calf serum, 2mM L-glutamine and 1mM sodium pyruvate; antibiotics were not included so as to allow early identification of bacterial contamination if present. Cells were grown at 37°C in a 5% CO_2 humidified atmosphere. The cells were split when confluent.

10.2   Cell splitting

The medium was removed from the cells in a 175 cm² flask and the monolayer washed twice with 5 mls warmed (37°C) versene (1:5000 EDTA in isotonic saline, Gibco-BRL Ltd, UK), then once with 5 mls warmed 1X trypsin (Sigma chemicals Ltd, UK). Approximately 1.5 mls of this trypsin was left on the cells until they detached (up to 10 mins). These cells were suspended in 10 mls of warmed growth medium by pipetting up and down around 30 times through a 10 ml pipette. To split 1:5, 2 mls of the suspension were mixed with 30 mls of medium in a new flask.

10.3   Freezing mammalian cells

The medium was removed from a 175 cm² flask of cells and the monolayer washed as in **10.2** except only 0.5 ml of trypsin was left behind. The detached cells were suspended in 9.5 mls of freezing medium (100 ml comprising DMEM 62 ml, DMSO 8 ml mixed and filter sterilized and foetal calf serum 30 ml). 1 ml aliquots were placed in 1.8 ml Nunc cryotubes and then placed in an insulated box at -70°C for 4 hours to gradually reduce the temperature before immersing in liquid nitrogen.

APPENDIX 10

Buffers, Media and Solutions

TAE 50x buffer
242g   Tris base
57.1ml   glacial acetic acid
100ml   0.5M EDTA, pH 8.0
Adjust to pH 7.2 and bring to final volume 1 litre with deionised water

TBE 10x buffer
107.8g   Tris base
55g   boric acid
7.44g   disodium EDTA, dihydrate
Add components in order above to 800ml deionised water. Add slightly less than the total amount of boric acid. When dissolved, adjust pH to 8.3 with remaining boric acid and bring final volume to 1 litre with deionised water.

Alkaline electrophoresis buffer 1x
5 ml 10N NaOH
2 ml 0.5M EDTA pH 8.0
Make up to 1 litre with deionised water

TE buffer
10mM   Tris-Cl, pH 8.0 (pH adjusted for different applications)
1mM   EDTA pH 8.0

TNE buffer
100mM   NaCl
10mM   Tris-Cl, pH 8.0
1mM   EDTA, pH 8.0

EDTA, 0.5M
93.05g   disodium ethylenediaminetetraacetate, dihydrate
Add the EDTA to 300ml deionised water, adjust to pH 8.0 with NaOH pellets. When dissolved, adjust the final volume to 500ml and autoclave.

PEG-NaCl
For 20% PEG, 2.5M NaCl dissolve 20g of PEG and 14.6g NaCl in 100 ml of deionised water.

3M sodium acetate pH 5.2
dissolve 408.1g sodium acetate (trihydrate) in 800 ml deionised water. Adjust pH with glacial acetic acid and make up volume to 1 litre.

SSC, 20x
175.3g   NaCl
88.2g   sodium citrate
In 800ml of deionised water, adjust to pH 7.2 with 10N NaOH, make up to 1 litre, autoclave as necessary. Use DEPC treated water for use with RNA.

SSPE, 20x
175.3g   NaCl
27.6g $Na_2HPO_4.H_2O$
7.4g   EDTA, disodium salt
Dissolve in 800ml of deionised water, adjust to pH 7.4 with 10N NaOH, make up to 1 litre, autoclave as necessary. Use DEPC treated water for use with RNA.

10% SDS
Dissolve 100g electrophoresis grade SDS in 900 ml deionised water. Heat to 68°C to assist dissolution. Adjust pH to 7.2 with concentrated HCl and make up to 1 litre with water. Wear mask to avoid breathing in SDS dust.

Denhardt's Reagent, 50x
5g   Ficoll (Pharmacia, type 400)
5g   polyvinylpyrrolidone
5g   bovine serum albumin (fractionV, Sigma)
Dissolve in deionised water (DEPC treated if for use with RNA) and make up to 500ml. Filter sterilise (0.45μm) and store at -20°C

BLOTTO- bovine lacto transfer technique optimizer
Southern blot blocking agent. 5% non-fat dried milk powder dissolved in deionised water. Use freshly prepared reagent each time.

DEPC treated water
Add diethyl pyrocarbonate (DEPC) to deionised water to a final concentration of 1%. Incubate overnight at room temperature in a fume hood. Autoclave. DEPC is a suspected carcinogen.

1M Tris pH 7.4-8.0
Dissolve 121.1g Tris base in 800 ml deionised water, adjust the pH with concentrated HCl as follows (make pH measurements at RT):
PH 7.4- 70 ml HCl; pH 7.6- 60 ml HCl; pH 8.0- 42 ml HCl
Make up to 1 litre with deionised water.

PBS
0.2g KCl
8.0g NaCl
0.2g KH_2PO_4
1.15g Na_2HPO_4
Add components to 900ml deionised sterile water. When completely dissolved, adjust to pH 7.4 with 1N HCl or 1N NaOH. Final volume 1 litre. Filter sterilise if storing long term.
For PBS-Tween, include 0.1% Tween-20.

RNA sample buffer
10.0ml   deionised formamide
3.5ml   37% formaldehyde
2.0ml   MOPS 5x buffer (0.2M MOPS pH 7.0, 50mM sodium acetate, 5mM
              EDTA pH 8.0)
Aliquot and store at -20°C. Formamide is a teratogen and formaldehyde is a toxic carcinogen.

RNA loading buffer
50%   glycerol
1mM   EDTA
0.4%   bromophenol blue
Use nuclease free water and store at -20°C.

Alkaline gel running buffer
50mN   NaOH
1mM   EDTA
Make fresh when required, use DEPC treated water

Tris-glycine buffer (SDS-PAGE)
25mM   Tris
250mM   glycine (electrophoresis grade)
0.1%   SDS (electrophoresis grade)

SDS sample buffer, 2x
100mM   Tris-HCl, pH 6.8
2% β-mercaptoethanol (other have DTT)
4%   SDS (electrophoresis grade)
0.2%   bromophenol blue
20%   glycerol

Western blot transfer buffer
4.545g   Tris base
54g   Glycine
Methanol   300ml
Made up to 1.5 litres using deionised water.

High salt RIP buffer
10mM   Tris-HCl pH 7.4
600mM   KCl
150mM   NaCl
0.5mM $MgCl_2$
1%   Triton X-100

PMSF (100mM stock in ethanol) is added to give a final concentration of 1mM immediately before use.

Phenol:chloroform:isoamyl alcohol (25:24:1)
Mix equal parts of TE buffer and phenol and allow the phases to separate. Then mix 1 part of the lower, phenol phase with 1 part of chloroform:isoamyl alcohol (24:1).

LB medium
10g tryptone
5g   yeast extract
5g   NaCl
Add deionised water to around 980ml. Adjust to pH 7.5 with 10N NaOH. Bring to 1 litre and autoclave.

LB agar
Add 15g of agar to 1 litre of LB medium and autoclave.
Addition of antibiotics should be delayed until the solution has cooled to 55°C or less.

LB top agar
Add 0.8g of agar to each 100ml LB medium and autoclave.

SOC medium
2.0g   tryptone
0.5g   yeast extract
1ml   1M NaCl
0.25ml   1M KCl
1ml   2M Mg^{2+} stock (1M $MgCl_2.6H_2O$, 1M $MgSO_4.7H_2O$), filter sterilised
1ml   2M glucose, filter sterilised
Dissolve tryptone, yeast extract, NaCl and KCl in 97ml deionised water and autoclave. Cool to room temperature and add magnesium stock and glucose. Filter medium through 0.2μm unit, pH should be 7.

APPENDIX 11

Table of potential genotypic designations for isolates of HEV

Isolate	Scheme 1[a]	Scheme 2[a]	Schlauder[b]	Wang[c]	Arankalle[d]	Tsarev[e]
B1, B2, I2	I	1	1	1a	IA	I1b
P1, C1-4	I	1	1	1b	IB	I1a
I1	I	1	1	1c	-	I1c
M1	II	2	2	2	II	II
Ni	II	-[f]	-	-	-	-
US1, US2, sUS	III	3	3	3	III	III
It1	III	4	5	-	-	-
sNZ	III	4	-	-	-	-
G1, Sp1, Sp2	III	5	6	-	-	-
G2	III	6	7	-	-	-
Ar1, Ar2	III	7	8	-	-	-
T1	IV	8	-	-	-	-

Table adapted from Schlauder and Mushahwar, 2001

Dash (-) indicates isolate not included in analysis.
[a]Schlauder and Mushahwar, 2001; [b]Schlauder et al., 2000; [c]Wang et al., 1999; [d]Arankalle et al., 1999b; [e]Tsarev et al., 1999.
[f]overlapping sequence from some isolates not available.
The Genbank accession numbers relating to the isolates are as follows:
(I1 is the study isolate)
Full length sequences:
I1- X98292 (Donati 1997); I2- X99441 (von Brunn, unpublished); I3- AF076239; B1-M73218 (Tam 1991); B2- D10330 (Aye 1993); P1- M80581 (Tsarev 1992); C1- D11092 (Aye 1992); C2- L25547 (Yin, unpublished); C3- M94177 (Bi 1993); C4- D11093 (Uchida, unpublished); M1- M74506 (Huang 1992); US1- AF060668 (Schlauder 1998); US2-AF060669 (Schlauder 1998); T1- AJ272108 (Wang 2000). sUS- swine US isolate AF082843 (Meng 1997); sNZ- New Zealand swine isolate AF200704 (Garkavenko 2000).
Partial length sequence:
Ni- AF172999 (Buisson 2000); It1- Italian isolate AF110387 (Schlauder 1999); G1- Greek isolate AF110388 (Schlauder 1999); G2- Greek isolate AF110389 (Schlauder 1999) Ar1-Argentinian isolate AF264009 (Schlauder 2000); Ar2- Argentinian isolate AF264010 (Schlauder 2000).

APPENDIX 12

Relationship between virus excretion, icterus/hepatitis and antibody production following human experimental HEV infection
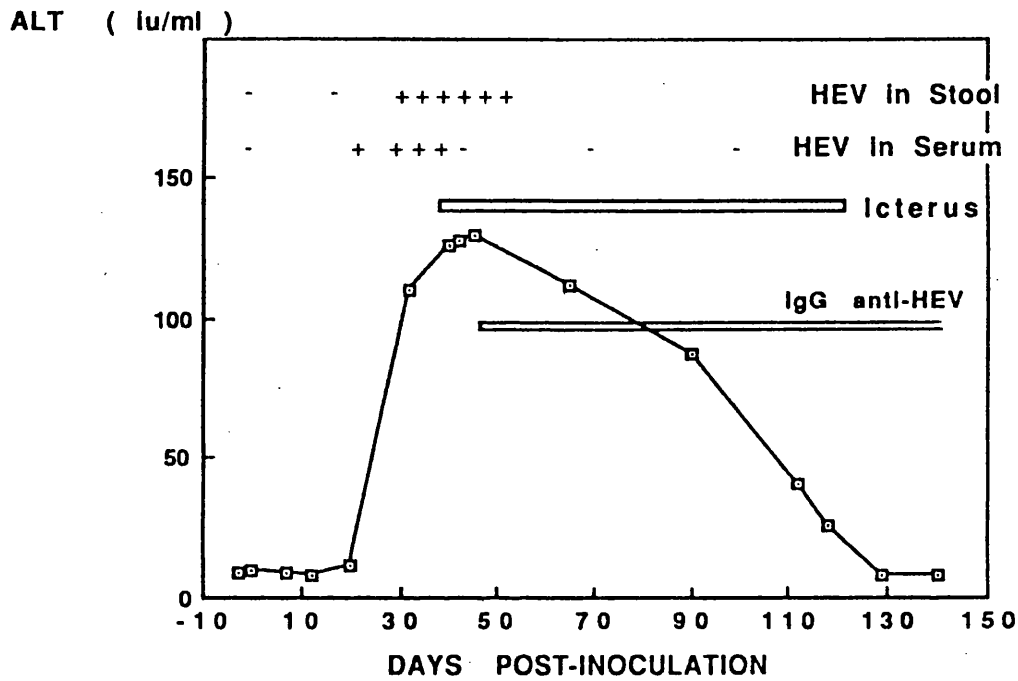

Figure taken from Chauhan *et al.,* 1993.
The presence of HEV in stool and serum relates to detection of HEV nucleic acid by RT-PCR.

Sequence analysis of full length HEV clones derived directly from human liver in fulminant hepatitis E

M.C. DONATI, E.A. FAGAN, T.J. HARRISON

University Department of Medicine
Royal Free Hospital School of Medicine, UK

Correspondence to: T.J. Harrison, University Department of Medicine
Royal Free Hospital School of Medicine, Rowland Hill Street
London NW3 2PF, UK

Abstract

The entire nucleotide sequence was determined for hepatitis E virus acquired in India and associated with a sporadic case of acute liver failure in a male. The genome was amplified by reverse transcription-polymerase chain reaction using human liver as the source of RNA and long range PCR to amplify the entire 7.2 kb cDNA. To our knowledge this is the first full HEV sequence to be generated directly from human liver and implicated directly in fulminant hepatitis.

The nucleotide, and predicted amino acid, sequences of this isolate were compared to others in the database, most of which are derived from virus passaged through monkeys. The closest homology (94% nucleotide, and 98% amino acid, identity) was with a Chinese isolate cloned from rhesus monkey bile. A partial sequence from India covering the non-structural region of HEV had 92% nucleotide, and 97% amino acid, identity whereas the most distantly related isolate (77 % nucleotide, 82% amino acid, identity) was that from Mexico. There were no obvious features of our HEV sequence which might account for the severe outcome of infection. The overall similarity in sequence between our non-passaged isolate and others after limited passage in non-human primates does not suggest significant adaptation to that host. Determination of the interplay between viral and host factors in the pathogenesis of fulminant hepatitis E will require comparative analyses with other HEV sequences, as well host factors, from defined cases covering the range from symptomless to fulminant infection in pregnancy.

Introduction

HEV is the major cause of enterically transmitted non-A non-B hepatitis[1] and is endemic throughout most of the developing world where it causes large epidemics as well as sporadic cases of acute viral hepatitis[2-5]. Reports of sporadic disease in Europe[6, 7] indicate that infections occasionally may occur in areas of low prevalence and in the absence of obvious risk factors such as travel to endemic areas.

Acute HEV infection can result in asymptomatic or clinically apparent disease but, although the viraemia may be prolonged[8], there is no evidence for chronicity or a carrier state. HEV is implicated as a cause of fulminant hepatitis in endemic areas where between 8%[9] and 40%[10] of cases have HEV as the only cause. A mortality rate of up to 20% has been reported for the hepatitis E in pregnancy[11].

Morphologically, HEV is similar to the Caliciviridae but formal classification is awaited. The genomes of several isolates of HEV

have been cloned and sequenced, with up to 24% variation in nucleotide sequence between isolates[12-15]. HEV has a positive stranded RNA genome of around 7.2kb (including a polyA tail) which contains three, conserved open reading frames. The amino-terminal ORF 1 is about 5kb in size and encodes the putative non-structural polyprotein which contains motifs characteristic of methyltransferase, papain-like protease, RNA helicase, and RNA-dependent RNA polymerase domains. Computer analysis of the domains in ORF 1 has led to the suggestion that HEV belongs to an alphavirus-like supergroup[16]. The carboxyl-terminal 2.2kb encodes the putative structural protein(s) in ORF 2. The small ORF 3 largely overlaps ORF 2 and also may encode a structural component of the virion.

The pathogenesis of hepatitis E is poorly understood, especially the relative contributions of host and viral factors to disease severity, including fulminant hepatitis in pregnancy. Recent attempts to reproduce fulminant infection in pregnant macaques have not been successful[17].

Most reported HEV sequences were derived from virus after passage in monkeys and it is not clear whether the virus has adapted to that host. In order to determine whether viral factors contribute to pathogenicity, we used explanted liver tissue from a male with sporadic, acute liver failure as the source of HEV RNA for cloning and sequencing. We developed an RT-PCR method which allowed amplification of the full-length HEV genome to facilitate this analysis and the subsequent expression of HEV gene products *in vitro*.

Materials and methods

Patient

A 57 year old male developed acute liver failure following travel in India. Serological tests for HAV, HBV and HCV, and PCR for the respective nucleic acids in liver, were negative. HEV RNA was detected by RT-PCR in the explanted liver[18] and serological findings, including high IgM anti-HEV, were diagnostic of recent hepatitis E[19].

RNA extraction and reverse transcription

Total RNA was extracted from explanted liver tissue using guanidinium isothiocyanate and ultracentrifugation through caesium chloride[18], aliquoted into 5μg lots and stored under ethanol at 70°C.

5μg of the this RNA was reverse transcribed into first strand cDNA in a 50μl reaction containing 5μg oligo-dT$_{12-18}$ as primer,

1000 units SuperscriptII reverse transcriptase (Life Technologies Ltd, UK), 0.5mM each dNTP, 130 units HPRNI (USB, USA) and first strand synthesis buffer. The reaction was carried out at 45 °C for 1 hour.

Long range PCR

1-2 µl of the cDNA synthesised in the above reaction were used in the PCR. The GeneAmp XL PCR kit (Perkin-Elmer, USA) was employed in a 100µl reaction containing 200µM each dNTP, 1.25mM magnesium acetate, 0.4µM primers (Tab. 1), buffer supplied by the manufacturer, and 3 units of rTth polymerase. A buffer giving final reaction concentrations of 25mM tricine/85mM potassium acetate (pH 8.7), 8% glycerol and 2% DMSO was found to be equally effective. Hot start was achieved using Ampliwax PCR Gems (Perkin-Elmer, USA) and thermal cycling was carried out in a Perkin-Elmer 9600 machine using the following parameters: an initial denaturation of 94°C 1 min, then 20 rounds of 94 °C 15 sec, 68 °C 7 min followed by 17 rounds of 94 °C 15 secs, 68 °C 7 min with autoextension at 15 sec per cycle and a final extension of 10 min at 72 °C. 2-10µl of this PCR reaction was amplified in a second, identical reaction to yield a product visible on agarose gel electrophoresis.

5' RACE

One aliquot of first strand cDNA from above was treated with RNase H and purified using silica beads (Qiaex II, Qiagen Ltd, UK), according to the manufacturers protocol, into a final volume of 40µl water. 16.5µl was used in each of two homopolymer tailing reactions: 25µM dATP or dGTP and 20 units terminal deoxynucleotidyl transferase (Promega Ltd, UK) using the buffer supplied by the manufacturer and incubation at 37°C for 30 min. The tailed cDNA was purified using silica beads, eluted into 40µl water and 1µl of each product was used in the following PCR: GeneAmp XL PCR kit (Perkin Elmer) buffer, 200µM dNTPs, 1.1mM magnesium acetate, 0.2µM primers (semi-nested, Tab. 1) and 2 units of rTth polymerase supplied in a 100µl reaction. Hot start cycling using Ampliwax PCR gems (Perkin Elmer) was carried out in a Perkin Elmer 9600 thermal cycler using the following parameters: initial denaturation 94°C 1 min, 37 rounds of 94 °C 15 sec, 50 °C 20 sec, 72 °C 2 min and a final extension of 10 min at 72 °C.

3' RACE

First strand cDNA was synthesised as above except that 50 pmoles of 3' RACE oligonucleotide (Tab. 1) was used as primer. 1µl of this

Table I. — Primers used for full-length PCR and to determine the sequences of the 5' and 3' ends of the genome.

5' posn 1-26*	AGGCAGACCACATATGTGGTCGATGC
3' posn 7194-7169	CAGGGAGCGCGGAACGCAGAAAAGAG
T primer (RACE)	GCTCTAGAAGCTTTTTTTTTTTTTTTTT
C primer (RACE)	GGAATTCGCGGCCGCCCCCCCCCCCCCCCCCCC
3' RACE GSP 1 6635-6654	gCgGattCCaTGGTTAATGT
3' RACE GSP 2 6971-6989	CTGCTTGAGGATACCATGG
5' RACE GSP 1 542-516	GATGGTGACATATCATGAAGGGAGTA
5' RACE GSP 2 478-457	AAAACCCGTCGAGGCAGTAAGT

*Numbering based on the Burmese prototype sequence (accession no. M73218). Lower case letters indicate mismatches with our consensus sequence.

cDNA was amplified in a 100µl PCR using 0.3µM primers (semi-nested, Tab. 1), 200µM dNTPs, 5% glycerol, 1.5mM magnesium chloride, 500mM KCl, 100mM Tris-HCl pH 8.3, and 2 units Taq DNA polymerase (Advanced Biotechnologies Ltd, UK). Thermal cycling in the PE 9600 was carried out with 94 °C 2 min then 38 rounds of 94 °C 1 min, 50 °C 30 sec, 72 °C 1 min 30 secs followed by a final extension of 10 min at 72 °C.

Cloning and sequencing

Full length HEV PCR products were cloned in two ways: the DNA was either blunt end ligated into the pCNTR vector (5 Prime 3 Prime inc, USA) or inserted into the pCRII TA cloning vector (Invitrogen BV, The Netherlands) following the addition of 3' dATP. Manual plasmid dideoxy sequencing was carried out using Sequenase version 2.0 (Amersham International plc, UK).

The consensus sequence for our isolate was generated from PCR products spanning the regions 4547-4925 and 6141-6870. the RACE products and completed using 3 full length clones (one from an independent PCR).

Analysis of sequences

The nucleotide and predicted protein sequences were compared to the EMBL and Genbank databases using BLAST[20]. The full length nucleotide sequence was aligned with the 10 full length sequences in the database using PILEUP and displayed using PRETTY (GCG software[21]).

Results

Our original strategy was to amplify sequences covering the entire genome in a number of semi-nested PCR reactions using primers synthesised on the basis of the prototype Burmese sequence. Although this approach generated around 2kb of sequence data, reactions utilising a number of primer sets did not yield detectable products. We sought, therefore, to amplify full-length cDNAs using long range PCR. This approach also was designed to obviate the need for assembling clones prior to expression.

Five full-length clones were derived, four following treatment of the full-length product with Taq polymerase and dATP to extend the 3' ends and permit cloning into the pCRII vector. Subclones for further sequence analysis were derived following Hind III and Eco RI digestion. The bulk of the sequence data was derived by plasmid sequencing using HEV-specific oligonucleotides as primers. The entire sequence, which has been submitted to the EMBL database (Accession No X98292), is derived from a consensus of at least 3 independent reactions.

The sequence was compared to those in the databases using BLAST[20] and aligned to 10 full length sequences using PILEUP and displayed using PRETTY (GCG Software[21]). Figure 1 shows the dendrogram used to create the PILEUP. Our sequence has around 94% nucleotide identity with several from Asia. The closest match is the sequence of a Chinese isolate from bile after the virus was passaged twice in macaques.[22] Most of the nucleotide variations in our sequence do not result in amino acid substitutions.

Table II lists the numbers of nucleotide substitutions by open reading frame (and in the untranslated regions) and those which are not found in any of the reported full length sequence. There are 11 amino acid substitutions in ORF 1 and 5 in ORF 2 which are not found in any other sequences in the database (Tab. III).
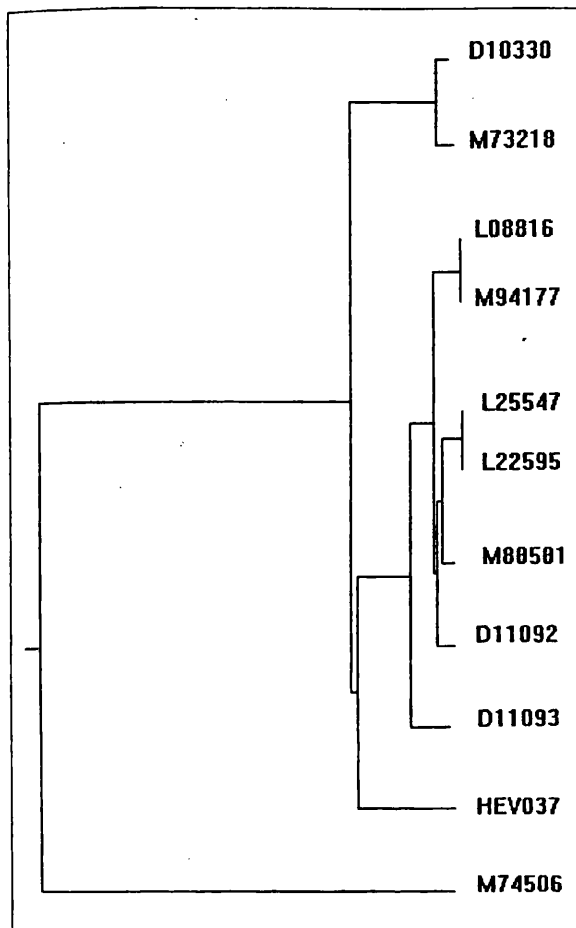
Fig. 1. — Dendrogram generated by PILEUP (21) of our Indian sequence and all other full length HEV sequences available in the databases. The horizontal axis represents degree of divergence: the Mexican isolate. M74506, has 77.5% similarity with ours, all the others are of Asian origin and show 93-94% similarity with this sequence from India. The sequence database accession numbers, with references and the source of virus, where known are: D10330, Burma27; M73218, Burma12; L25595, China22; M80581, Pakistan14 and D11092, China15. All the above are derived from virus passaged in monkeys. L08816, China26 was isolated directly from human stool. M74506, Mexico13 was derived partly from virus passaged in monkeys and partly from virus in human stool. HEV037, India was generated in this study directly from human liver as detailed and is the only sequence known to be from a case of acute liver failure. M94177, L25547 and D11093 are unpublished.

TABLE II. — *Nucleotide differences in various regions of the genome.*

Genomic Region	Nucleotide Substitutions vs Closest Match (L25595)	Unique Nucleotide Substitutions
5' UTR	1*	1*
ORF 1	306	98
ORF 2	120	36
ORF 3	11	3
3' UTR	2	0

Column 2 - versus the closest match in the database (Chinese isolate)
Column 3 - unique nucleotides not found in other reported sequences
*A to C transversion at nucleotide position 11.

TABLE III. — *Unique amino acid substitutions. Square brackets are those predicted from other sequences. The sequence of one of our three full length clones predicts the identical residue to that in the other Asian sequences.*

Genomic region	Details of unique amino acids substitutions
ORF 1	139 Ser [Pro]
	663 Ala [Ser]
	722 Ala [Thr, Leu, Ile, Met]
	740 Gln [Leu, Val]
	748 Phe [Ser, Pro]
	763 Val [Ala, Ser]
	766 Val [Gly, Arg]
	827 Thr [Ala, Asp]
	928 Ala [Thr]
	938 Ala [Val]
	1225 Thr [Ala, Arg]
ORF 2	55 His [Tyr]*
	97 Thr [Ala]
	146 Pro [Ser]*
	283 Leu [Ser]*
	306 Phe [Leu]*

Discussion

We amplified using RT-PCR a cDNA copy of the entire genome of HEV, contracted in India and associated with a sporadic case of acute liver failure, using RNA extracted from explanted liver. Analysis of our isolate shows conservation of the genomic organization of the other published HEV sequences.

The 5' UTR of our Indian isolate is 26 bases in length, but is missing the initial adenine seen in the Burmese and some Chinese sequences. There is also an A to C tranversion at position 11 which falls within the loop of the proposed stem-loop structure formed by the first 27 bases[13]. This nucleotide substition is not seen in any of the other available HEV sequences and is in a region that may influence viral expression. so may be significant. The 3' UTR sequence of our isolate also is highly conserved compared to published sequences and terminates in poly A (at least 23 bases).

The first open reading frame shows greatest divergence from sequences in the database, particularly in the hypervariable region (at around nucleotides 2000-2400) where 5 of the 9 unique amino acids substitutions are situated. All of the expected motifs within the putative ORF 1[16] were conserved in our isolate. These include the residues CTC, postulated as the active site in papain-like cysteine proteases, the NTP-binding domains GVPGSGKS and DEAP in the putative helicase region and the GDD site found in all RNA dependent RNA polymerases.

The putative structural gene of ORF 2 is most variable in the N-terminal region with all 5 unique amino acids (Tab. III) occurring before the defined epitopes, determined previously to be sited after residue 394[23, 24]. As noted in Table III, four of these five residues were not variant in one of our sequences. This indicates heterogeneity which may be the result of antibody pressure. There are 19 amino acid changes in ORF 2 compared to another sequence of Indian origin[25].

The predicted 123 residue ORF 3 protein has two substitutions (leucine for serine at residue 76 and alanine for valine at residue 89) compared to the closest matches. Serine and alanine are conserved at these positions in all other Asian isolates in the databases. However, the Mexican strain[13], which is most divergent (14/124 amino acid substitutions), has leucine and alanine at these positions. There were 7 amino acid substitutions compared to another Indian isolate[25] which is 92% homologous at the nucleotide and 97% at the amino acid level over the ORF 3/2 region.

This Indian HEV sequence shows closest homology to other Asian isolates from geographically related areas but is closer over the structural region to Chinese and Burmese sequences than to another Indian isolate[25]. The sequence is most distant to that from Mexico but shares with that strain 9 amino acids (5 in ORF 1, 2 each in ORFs 2 and 3) which are not found in other sequences of Asian origin.

There were no obvious features of our HEV sequence which might account for the severe outcome of infection. The overall similarities in sequence between our non-passaged isolate from human liver, another derived directly from human stool[26] and those after limited passage, suggests no significant adaptation to non-human primates. Determination of the interplay between viral and host factors in the pathogenesis of fulminant hepatitis E will require comparative analyses with other HEV sequences, as well host factors, from defined cases covering the range from symptomless to fulminant infection in pregnancy. Expression of these clones *in vitro* is in progress.

Acknowledgements

References

1. Bradley DW. Hepatitis E: Epidemiology, aetiology and molecular biology. Rev. Med. Virol. 1992;2:19-28.
2. Khuroo MS. Study of an epidemic of non-A, non-B hepatitis. Possibility of another human hepatitis virus distinct from post-transfusion non-A, non-B type. Am. J. Med. 1980;68:818-824.
3. Skidmore SJ, Yarbough PO, Gabor KA, Reyes GR. Hepatitis E virus: the cause of a waterbourne hepatitis outbreak. J. Med. Virol. 1992;37:58-60.
4. Favorov MO, Fields HA, Purdy MA, Yashina TL, Aleksandrov AG, Alter MJ, Yarasheva, DM, et al. Serologic identification of hepatitis E virus infections in epidemic and endemic settings. J. Med. Virol. 1992;36:246-250.
5. Ticehurst J, Popkin TJ, Bryan JP, Innis BL, Duncan JF, Ahmed A, Malik I, et al. Association of hepatitis E virus with an outbreak of hepatitis in Pakistan: serologic responses and pattern of virus excretion. J. Med. Virol. 1992;36:84-92.
6. Zaaijer HL, Kok M, Lelie PN, Timmerman RJ, Chau K, van der Pal HJ. Hepatitis E in The Netherlands: imported and endemic. Lancet 1993;341:826.
7. Jardi R, Buti M, Rodriguez-Frias F, and Esteban R. Hepatitis E infection in acute sporadic hepatitis in Spain. Lancet 1993;341:1355-1356.
8. Nanda SK, Ansari IH, Acharya SK, Jameel S, Panda SK. Protracted viremia during acute sporadic hepatitis E virus infection. Gastroenterology 1995;108:225-230.
9. Khuroo MS, Rustgi VK, Dawson G J, Mushahwar IK, Yattoo GN, Kamili S, Khan BA. Spectrum of hepatitis E virus infection in India. J. Med. Virol. 1994;43:281-286.

10. Nanda SK, Yalcinkaya K, Panigrahi AK, Acharya SK, Jameel S, Panda SK. Etiological role of hepatitis E virus in sporadic fulminant hepatitis. J. Med. Virol. 1994;42:133-137.
11. Khuroo MS, Teli MR, Skidmore S, Sofi MA, Khuroo MI. Incidence and severity of viral hepatitis in pregnancy. Am. J. Med. 1981;70:252-255.
12. Tam AW, Smith MM, Guerra ME, Huang CC, Bradley DW, Fry KE, Reyes GR. Hepatitis E Virus (HEV) - Molecular Cloning and Sequencing of the Full-Length Viral Genome. Virology 1991;185:120-131.
13. Huang CC, Nguyen D, Fernandez J, Yun KY, Fry KE, Bradley DW, Tam AW, et al. Molecular cloning and sequencing of the Mexico isolate of hepatitis E virus (HEV). Virology 1992;191:550-558.
14. Tsarev SA, Emerson SU, Reyes GR, Tsareva TS, Legters LJ, Malik IA, Iqbal M, et al. Characterization of a prototype strain of hepatitis E Virus. Proc. Natl. Acad. Sci. USA. 1992;89:559-563.
15. Aye TT, Uchida T, Ma XZ, Iida F, Shikata T, Zhuang H, Win KM. Complete nucleotide sequence of a hepatitis E virus isolated from the Xinjiang epidemic (1986-1988) of China. Nucleic Acids Res. 1992;20:3512.
16. Koonin EV, Gorbalenya AE, Purdy MA, Rozanov MN, Reyes GR, Bradley DW. Computer-assisted assignment of functional domains in the nonstructural polyprotein of hepatitis E virus - delineation of an additional group of positive-strand RNA plant and animal viruses. Proc. Natl. Acad. Sci. USA. 1992;89:8259-8263.
17. Tsarev SA, Tsareva TS, Emerson SU, Rippy MK, Zack P, Shapiro M, Purcell RH. Experimental hepatitis E in pregnant rhesus monkeys: failure to transmit hepatitis E virus (HEV) to offspring and evidence of naturally acquired antibodies to HEV. J. Infect. Dis. 1995;172:31-37.
18. Fagan EA, Harrison TJ. Candidate hepatitis F virus in sporadic non-A, non-B acute liver failure: exclusion in liver of hepatitis viruses A, E, C and B by polymerase chain reaction. In: Nishioka K, Suzuki H, Mishiro S, Oda T, eds. Viral Hepatitis and Liver Disease. Tokyo: Springer-Verlag. 1994:73-76.
19. Fagan EA, Menon T, Valliammai T, Paul DA, Donati MC, Burroughs AK, Gunamanam D, et al. Equivocal serological diagnosis of sporadic fulminant hepatitis E in pregnant Indians. Lancet 1994;344:342-343.
20. Karlin S, Altschul SF. Methods for assessing the statistical significance of molecular sequence features by using general scoring schemes. Proc. Natl. Acad. Sci. USA. 1990;87:2264-2268.
21. Devereux J, Haeberli P, Smithies O. A comprehensive set of sequence analysis programs for the VAX. Nucleic Acids Res. 1984;12:387-395.
22. Yin SR, Purcell RH, Emerson SU. A new Chinese isolate of hepatitis E virus: comparison with strains recovered from different geographical regions. Virus Genes 1994;9:23-32.
23. Yarbough PO, Tam AW, Fry KE, Krawczynski K, McCaustland KA, Bradley DW, Reyes GR. Hepatitis E virus: identification of type-common epitopes. J. Virol. 1991;65:5790-5797.
24. Khudyakov YE, Favorov MO, Jue DL, Hine TK, Fields HA. Immunodominant antigenic regions in a structural protein of the hepatitis E virus. Virology 1994;198:390-393.
25. Jameel S, Zafrullah M, Ozdener MH, Panda SK. Expression in animal cells and characterization of the hepatitis E virus structural proteins. J. Virol. 1996;70:207-216.
26. Bi SL, Purdy MA, McCaustland KA, Margolis HS, and Bradley DW. The sequence of hepatitis E virus isolated directly from a single source during an outbreak in China. Virus Res. 1993;28:233-246.
27. Aye TT, Uchida T, Ma X, Iida F, Shikata T, Ichikawa M, Rikihisa T, et al. Sequence and gene structure of the hepatitis E virus isolated from Myanmar. Virus Genes 1993;7:95-110.